ELSEVIER
Available at
www.ElsevierComputerScience.com
POWERED BY SCIENCE DIRECT
Journal of Complexity 21 (2005) 111–148
Journal of COMPLEXITY
http://www.elsevier.com/locate/jco

Semialgebraic complexity of functions

Y. Yomdin

Department of Mathematics, The Weizmann Institute of Science, Ziskind 153, P.O. Box 26, 76100 Rehovat, Israel

Received 22 February 2003; revised 15 September 2003; accepted 16 September 2003

Abstract

In this paper we study the rate of the best approximation of a given function by semialgebraic functions of a prescribed "combinatorial complexity". We call this rate a "Semialgebraic Complexity" of the approximated function. By the classical Approximation Theory, the rate of a *polynomial* approximation is determined by the regularity of the approximated function (the number of its continuous derivatives, the domain of analyticity, etc.). In contrast, semialgebraic complexity (being always bounded from above in terms of regularity) may be small for functions not regular in the usual sense. We give various natural examples of functions of low semialgebraic complexity, including maxima of smooth families, compositions, series of a special form, etc. We show that certain important characteristics of the functions, in particular, the geometry of their critical values (Morse–Sard Theorem) are determined by their semialgebraic complexity, and not by their regularity.

Keywords: Complexity; Semialgebraic approximation; Critical points and values

1. Introduction

This paper summarizes a part of the results obtained during the last 15 years via a certain specific approach to a nonlinear approximation, representation, and processing of analytic and empiric data.

Our main conclusion can be informally summarized as follows: *Complexity of many types of mathematical objects may be much lower than their "regularity" suggests. A nonlinear approximation by the adequate high order approximants (like*

E-mail address: yosef.yomdin@weizmann.ac.il.

0885-064X/$ - see front matter
doi:10.1016/j.jco.2003.09.005

semialgebraic functions), *taking into account the geometry of singularities of the object*, *detects and utilizes this low complexity*, *both in storage and in processing*.

This paper deals with the "analytic" part of this general approach. Its main result is that *the rate of the best approximation of a given function by semialgebraic functions of a prescribed* "*combinatorial complexity*" (*we call below this rate the* "*semialgebraic complexity*" *of the approximated function*) *determines some important analytic characteristics of the function*, *which have been traditionally associated with its regularity*.

In this paper we use the validity of the Morse–Sard Theorem (see below) as the analytic (geometric) property to be traced. We mention only very shortly other important analytic properties which can be investigated in the framework of semialgebraic complexity: geometry of the level sets, behavior under iterations, etc.

This paper can be considered as a continuation of the paper [Yom18]: "Complexity of Functions: some questions, conjectures and results" by the author, published in the Journal of Complexity in 1991. First of all, we prove some of the results stated without proofs in [Yom18]. This concerns maximum functions, lacunary series, and compositions. We prove also several new results in these directions, as well as in a new direction of "fewnomial complexity", introduced here. The last section is devoted to the semialgebraic complexity of functions on infinite-dimensional spaces. We state and discuss results from [Yom17] estimating complexity of certain classes of functions on ℓ^2 and on $C^k[0,1]$. Then we apply results of [Bri-Yom1,Bri-Yom2,Bri-Yom3,Bri-Yom4] in order to bound a semi-algebraic complexity of the input-to-state mapping of certain control problems.

The proposed approach has been experimentally tested in several directions. We plan to present separately a description of the implemented algorithms as well as the experimental results. Here we just mention shortly the main investigated problems.

The Taylor piecewise-polynomial approximation has been used for solving elliptic and parabolic PDE's in [Bic-Yom,Ko-Se-Yom1,Ko-Se-Yom2,Ko-Se-Yom3,Wie-Yom]. We have achieved a very high order of the discretization error, as compared to the size of the processed data, while preserving a stability of the classical methods. The investigation of singularities (like in Burgers equation) has been started.

Motion planning in Robotics has been investigated in [Eli-Yom1,Eli-Yom2,Eli-Yom3,Tan-Yom]. The piecewise-linear approximation has been used in this problem, while the main effort concentrated on the efficient treatment of singularities of the data (on the boundary of the "configuration space"). Efficient algorithms for a plane motion planning have been implemented.

An algorithm for a numerical inversion of nonlinear mappings, described in [Eli-Yom4], uses the Taylor piecewise-polynomial approximation together with the Whitney Normal forms of singularities ([Whi2], see also [Gol-Gui]) for a coherent representation of both the direct and the inverse mappings of the real plane into itself.

However, the most advanced implementation of our approach till now is in Image Processing. A nonlinear high-order approximation, together with the appropriate "Normal forms" capturing image singularities, has been used for a very compact representation of images and "synthetic video" sequences

[Bri-Eli-Yom,Eli-Yom5,Y-E-B-S]. The results have been intensively tested against various known methods, and have been used in working practical applications.

We believe that the implementation results till now confirm the efficiency of the approach partly presented in this paper.

Let us conclude with mentioning two basic (to our opinion) problems, penetrating all the area of the mathematical investigation of the complexity of approximation and processing of functions:

1. *What are natural measures of the approximation accuracy*? For example, a slightly "shifted" function (image) in most of applications should be considered as a very good approximation of the original one. On the other hand, for a function with singularities (like edges on the image) any reasonable norm of the difference of this function with its shift will be big.

2. *What are the advantages and the disadvantages of the linear versus nonlinear approximation and processing methods*? For example, images containing only black and white regions, separated by smooth edges, allow for a very compact nonlinear representation. Indeed, it is enough to memorize the geometry of the separating edges. On the other hand, any linear scheme for representing these images must be rather complicated.

We plan to present initial rigorous results in these directions separately.

2. Semialgebraic complexity

The classical result of Approximation Theory (see [Achi,Lor1,Lor2]) is that the rate of a polynomial approximation of a function is determined by its regularity, and vice versa. In "Constructive Function Theory" developed by Bernstein (see [Ber1,Ber2,Ber3] and many other classical papers and books by Bernstein) various analytic and metric properties of the approximated functions are obtained via the corresponding properties of the approximating polynomials.

In [Yom1,Yom2,Yom3,Yom4,Yom5,Yom6,Yom7,Yom8,Yom13,Yom14,Yom15,Yom16,Yom17,Yom18] it was shown that this approach can be extended to cover also geometric and topological properties of the levels sets of the approximated functions, of their critical points and critical values, etc. Here the "metric semialgebraic geometry" of the approximating polynomials is imposed onto the approximated functions. In [Yom-Com] this approach has been further extended: a wide class of metric and topological properties of semialgebraic sets and mappings has been investigated, which are "stable under approximation". It was shown in [Yom-Com] how these properties can be imposed onto smooth functions and mappings, via their global polynomial approximation or, alternatively, via their piecewise-polynomial approximation by Taylor polynomials.

It turns out, that exactly the same method can be applied to a much wider class of functions, than C^k-ones. In fact, what we need for the results of [Yom1,Yom2,Yom3,Yom4,Yom5,Yom6,Yom7,Yom8,Yom-Com] to be true, is not a regularity (C^k, C^ω, etc.) of the functions considered, but rather their "complexity", measured as the rate of their best approximation by semialgebraic functions of a given "combinatorial complexity". We call this rate a "semialgebraic complexity". (The situation is more delicate with the dynamical results of [Yom13,Yom14, Yom15,Yom16]. It turns out that these results cannot be extended directly to the functions of a low semialgebraic complexity.)

In this paper we present some definitions and results on semialgebraic complexity, stressing simple illustrating examples, as well as open problems and promising investigation directions.

The results on the fewnomial complexity, as well as a part of the results on the complexity of compositions, of lacunary series, and of control problems, are new. The other have been published in [Yom16,Yom17,Yom18].

To simplify a presentation, we always assume our functions to be continuously differentiable, and as the main property under investigation we take the one given by the Quantitative Sard Theorem ([Yom1], Theorem 3.1 below). Both these restrictions are not essential. In fact, on one side, the approach can be generalized to Lipschitzian functions (see [Cla1,Cla2,Yom9,Yom10]), and on the other side, much more general "geometric complexity bounds" (in the spirit of [Yom3]) can be obtained.

2.1. Definition of Semialgebraic complexity

Let $f: B^n_r \to \mathbb{R}$ be a C^1-function. Here B^n_r is the closed ball of radius r in $\mathbb{R}^n$. Let $g: B^n_r \to \mathbb{R}$ be a semialgebraic function (i.e. the graph of g is a semialgebraic subset of $\mathbb{R}^{n+1}$, defined by a finite number of polynomial equations and inequalities and set-theoretic operations). We do not assume g to be differentiable or even continuous. However, g is analytic on a complement of a semialgebraic subset $S(g)$, $\dim S(g) < n$.

Definition 2.1. For f, g as above, the deviation $|f-g|_{C^1}$ is defined as

$$|f-g|_{C^1} = \sup_{x \in B^n_r \setminus S(g)} (|f(x)-g(x)| + r\|\nabla f(x) - \nabla g(x)\|),$$

where $\| \ \|$ is the usual Euclidean norm of the gradients.

Now assume that a certain "complexity measure" C is given, that associates to each semialgebraic function $g: B^n_r \to \mathbb{R}$ a positive number $C(g)$, interpreted as a complexity of g. We do not consider $C(g)$ as *the* complexity of this semialgebraic function. In contrary, the choice of $C(g)$ is one of the important instruments in our approach: it reflects the specifics of the problem considered.

In most of examples $C(g)$ satisfies certain natural requirements, which we do not discuss explicitly, since below we use only one "complexity measure" C, specially adopted to the Quantitative Sard Theorem. However, it is important that for any explicitly given semialgebraic function g its complexity $C(g)$ be explicitly bounded from above in terms of the "combinatorial data" of the representation of g. This includes the degrees of the polynomials involved and the set-theoretic formula of the representation. We call below these data a diagram of g, $D(g)$.

Definition 2.2. Let $f: B_r^n \to \mathbb{R}$ be a C^1-function. A C-semialgebraic complexity $\sigma_s(f, \varepsilon)$, for any $\varepsilon > 0$, is defined as follows:

$$\sigma_s(f, \varepsilon) = \inf C(g),$$

where infimum is taken over all the semialgebraic functions g, such that $|f - g|_{C^1} \leqslant \varepsilon$.

In other words, $\sigma_s(f, \varepsilon)$ is the minimal "$C(g)$-complexity" of semialgebraic functions g, ε-approximating f in C^1-norm. Alternatively, we can define a "C-semialgebraic approximation rate" $E_s(f, d)$ as the $\inf|f - g|_{C^1}$ over all the semialgebraic g with $C(g) \leqslant d^n$.

These definitions are motivated by the classical results of approximation theory, where g are mostly taken to be polynomials (trigonometric polynomials, other orthogonal systems, etc.), and $C(g)$ is d^n for d the degree and n the number of variables. One of the most basic facts here is that the rate of a polynomial approximation of a given function is completely determined by its "regularity" in the usual sense: the number of continuous derivatives, in the finite smoothness case, or the size of the complex domain to which the function can be extended, in the real analytic case.

More accurately, let us assume that the "complexity measure" $C(g)$ satisfies the following additional requirement: $C(p) = d^n$ for any polynomial p of degree d. (C below, as well as complexity measures arising in other natural examples, satisfies this up to certain constants.)

Define the polynomial "complexity" and "approximation rate" as

$$\sigma_{\text{poly}}(f, \varepsilon) = \inf_p \ C(p),$$

over all polynomials p with $|f - p|_{C_1} \leqslant \varepsilon$,

$$E_{\text{poly}}(f, d) = \inf_p \ |f - p|_{C^1},$$

over all polynomials p with $C(p) \leqslant d^n$ (in other words, with the degree of p at most d) Written as above, the definition shows that σ_{poly} and E_{poly} are constructed exactly as σ_s and E_s, only with a subclass of all semialgebraic functions. This proves immediately that for any $\varepsilon > 0$ and $d > 0$,

$$\sigma_s(f, \varepsilon) \leqslant \sigma_{\text{poly}}(f, \varepsilon),$$

$$E_s(f, d) \leqslant E_{\text{poly}}(f, d).$$

Now the classical Jackson's and Bernstein's theorems in approximation theory can be reformulated in our case (not completely accurately) as follows (see, for example, [Achi,Lor1,Lor2]):

Theorem 2.3. *If $f: B_r^n \to \mathbb{R}$ is C^k, then*

$$\sigma_{\text{poly}}(f, \varepsilon) \leqslant K_1 \left(\frac{1}{\varepsilon}\right)^{\frac{n}{k-1}},$$

$$E_{\text{poly}}(f,d) \leqslant K_2 \left(\frac{1}{d}\right)^{k-1}.$$

Conversely, if $\sigma_{\text{poly}}(f,\varepsilon)$ *increases slower than* $K\left(\frac{1}{\varepsilon}\right)^{\frac{n}{k-1}}$ *(or, equivalently,* $E_{\text{poly}}(f,d)$ *decreases faster than* $K'\left(\frac{1}{d}\right)^{k-1}$*), then f is k times continuously differentiable on* B_r^n*.*

For analytic functions the corresponding result is true, with

$$\sigma_{\text{poly}}(f,\varepsilon) \sim |\log \varepsilon|^n,$$

$$E_{\text{poly}}(f,d) \sim q^d, \quad q<1.$$

Let us return to semialgebraic complexity. As we have shown, it is bounded from above by the polynomial complexity. In fact, for generic C^k or analytic functions σ_s and σ_{poly} are equivalent. This can be shown by the "massiveness" arguments: the ε-entropy (see below) of the set of uniformly bounded C^k-functions with a low semialgebraic complexity (considered as a subset in C^0) is smaller than the ε-entropy of all the C^k-functions. Hence, most of them cannot have too small σ_s. The bound for the ε-entropy of "simple" functions follows, in turn, from the corresponding bound for semialgebraic sets. We plan to present detailed results in this direction separately.

On the other hand, we can see immediately, that the semialgebraic complexity can be small for functions, not regular in the usual sense. Indeed, let $f(x)$ be defined as

$$f(x) = \begin{cases} 0, & -1 \leqslant x \leqslant 0, \\ x^2, & 0 \leqslant x \leqslant 1. \end{cases}$$

f is C^1, but not C^2 on $[-1,1]$, and since f is itself semialgebraic, $\sigma_s(f,\varepsilon) \leqslant \text{const}$, and $E_s(f,d)=0$ for d big enough. The same is true for any (C^1) semialgebraic function f.

Below we give many examples of functions, whose semialgebraic complexity is better than their regularity prescribes. Then the following problem becomes a central one for understanding the relationship between "regularity" and "complexity properties":

Does low semialgebraic complexity imply existence of the high-order derivatives in a certain generalized sense?

There are some partial results in this direction. A very important class of nonsmooth functions with a low semialgebraic complexity is given by the maxima of smooth families (see Section 4). For a function f, representable as a pointwise supremum of a bounded in C^2-norm family of C^2-smooth functions, its generalized Laplacian $\tilde{\Delta} f$ (in a distribution sense) is shown in [Yom7] to be a measure with an explicitly bounded variation and singular part.

Question. *Is a similar property true for f with* $\sigma_s(f,\varepsilon) \sim (1/\varepsilon)^{n/2}$?

Another approach here is the following: A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to have a kth Peano differential at $x_0 \in \mathbb{R}^n$, if there exists a polynomial $P : \mathbb{R}^n \to \mathbb{R}$ of degree k, such that $|f(x) - P(x)| = o(||x - x_0||^k)$. The classical result in convex geometry—Alexandrov-Fenchel theorem—claims that any convex function has the second Peano differential almost everywhere (see [Al,AZ,BF]). Suprema of C^2-families can be shown to admit a representation as a difference of two convex functions (see [Roc,Sha-Yom,Yom6]) and hence are almost everywhere twice Peano differentiable.

Conjecture. *If the semialgebraic complexity of a* C^1*-function* $f : B_r^n \to \mathbb{R}$ *satisfies* $\sigma_s(f, \varepsilon) \leqslant K\left(\frac{1}{\varepsilon}\right)^{\frac{n}{k}}$, *then* f *has a kth Peano differential almost everywhere.*

This conjecture is strongly supported by the following result of Dolgenko ([Dol1,Dol2], see also [Iva1,Iva2]):

Define the *rational* complexity $\sigma_r(f, \varepsilon)$ of f exactly as in Definition 2.2, but restricting the approximating functions to the rational ones. Then, as shown in [Dol1,Dol2], the condition $\sigma_r(f, \varepsilon) \leqslant K\left(\frac{1}{\varepsilon}\right)^{n/k}$ does not imply even C^2-smoothness of f. However, the results of [Dol1,Dol2] state that under this condition f has a kth Peano differential almost everywhere.

The proof in [Dol1,Dol2] seems to apply to semialgebraic complexity with no essential modifications.

Thus for a "C^k-type" behavior of the complexity ($\sigma(f, \varepsilon) \sim \left(\frac{1}{\varepsilon}\right)^{n/k}$) we have some partial results and (hopefully) reasonable conjectures. The following interesting problem then naturally arises:

What kind of "regularity" can be expected for functions with an "analytic-type" behavior of complexity $(\sigma(f, \varepsilon) \sim |\log \varepsilon|^n)$?

In particular, we can expect for such functions existence, at almost every point of Peano differentials of any order. (By Dolgenko's results this is so for functions with low rational complexity σ_r.)

Do the Taylor series, defined in this way, converge? What is their relation with the original function?

3. Semialgebraic complexity and Sard theorem

The main result of this section is that the geometry of the set of critical values of the C^1-function is determined by its semialgebraic complexity (and not by its regularity, as it appears in standard settings of Sard-like results). So let $f : B_r^n \to \mathbb{R}$ be a C^1-function, $\Sigma(f)$ the set of its critical points, and $\Delta(f) = f(\Sigma(f))$ the set of its critical values.

The classical Sard (or Morse–Sard) theorem ([Mor,Sar1,Sar2]) claims that *the Lebesgue measure of* $\Delta(f)$ *is zero, if* f *is k times continuously differentiable, with* $k \geqslant n$.

Now in order to define a complexity measure C appropriate for the study of critical points and values, we first have to make all the relevant notions "stable under approximation''. This was done in [Yom1]: let $f : B^n_r \to \mathbb{R}$ be a C^1-function as above. For any nonnegative ε, we say that a point x in B^n is an ε-critical point of f if the norm of the gradient of f at x is at most ε. The value of f at an ε-critical point is called an ε-critical value of f. We denote the set of ε-critical points of f by $\Sigma(f, \varepsilon)$, and the set of ε-critical values by $\Delta(f, \varepsilon)$.

As far as the *conclusion* of the classical Sard theorem is concerned, it also must be strengthened. Indeed, the Lebesgue measure of a set may be zero, while its ε neighborhood may cover an interval, and consequently this notion is not well adapted to work with approximations. Following [Yom1], we use instead of the Lebesgue measure the ε-entropy in order to measure the ''massiveness'' of the critical values, or, in general, of relatively compact subsets of metric spaces. Let X be a relatively compact subset of a metric space Y. For any positive ε, the ε-entropy $M(\varepsilon, X)$ is defined as the smallest number of ε-balls in Y, covering X. Various properties of ε-entropy can be found, for example, in [Kol-Tih,Tih,Tri,War,Yom-Com].

The following result (a ''Quantitative Sard Theorem'') has been proved in [Yom1]:

Theorem 3.1. *Let $f : B^n_r \to \mathbb{R}$ be a C^k-function. Then for positive γ and ε,*

$$M(\varepsilon, \Delta(f, \gamma)) \leqslant K_1 + K_2\left(\frac{1}{\varepsilon}\right)^{\frac{n}{k}} + K_3\gamma\left(\frac{1}{\varepsilon}\right)^{\frac{n+k-1}{k}}.$$

The proof of this result is based on a Taylor piecewise-polynomial approximation of f and on the following ''polynomial'' version of the Quantitative Sard Theorem:

Theorem 3.2. *Let $p : B^n_r \to \mathbb{R}$ be a polynomial of degree d. Then for positive γ and ε,*

$$M(\varepsilon, \Delta(p, \gamma)) \leqslant K_4 + K_5\left(\frac{\gamma r}{\varepsilon}\right).$$

In particular, for $\gamma r = \varepsilon$,

$$M(\varepsilon, \Delta(p, \varepsilon/r)) \leqslant K_4 + K_5.$$

The constants K_4 and K_5 depend only on n and d and are of order of d^n.

The corresponding notions for a semialgebraic function g (not assumed to be smooth or even continuous) are defined in the same way, restricting to the complement of a semialgebraic subset $S(g)$, $\dim S(g) < n$, on which g is analytic.

Definition 3.3. For a semialgebraic function $g : B^n_r \to \mathbb{R}$ its ''Sard complexity'' $C_s(g)$ (or simply $C(g)$) is 3 times the supremum with respect to ε of the ε-entropy of the set $\Delta(f, \varepsilon/r)$ of ε/r-critical values of G.

The following generalization of Theorem 3.2 was obtained in [Yom5,Yom-Com]:

Theorem 3.4. *$C_s(g)$ is explicitly bounded in terms of the diagram $D(g)$.*

We do not give here this explicit expression (which is rather cumbersome and does not help much in specific examples). An important special case used below is the following:

Theorem 3.5. *For a piecewise polynomial g on a regular cell partition of the unit cell $[0,1]^n$ the bound for $C_s(g)$ is a constant multiplied by the sum of d_i^n, where d_i are the degrees of the participating polynomials. More generally, for a piecewise-polynomial function on a semialgebraic partition of the domain, its complexity $C_s(g)$ is bounded by the sum of d_i^n multiplied by the complexity of the corresponding semialgebraic piece.*

A regular cell partition of the unit cell $[0,1]^n$ is its subdivision into equal sub-cells of the size $\frac{1}{l}$. We do not give here an accurate statement of the second part of Theorem 3.5, in particular, an accurate definition of the complexity of a semialgebraic piece. Basically, it is equal to the chosen "Complexity measure" C_s applied to the boundary of the piece.

Now we finally can prove the main result of this section. Let as above $\sigma_s(f,\varepsilon)$ be the semialgebraic complexity of f, defined with respect to the complexity measure C_s.

Theorem 3.6. *For any $\varepsilon > 0$,*

$$M(\varepsilon, \Delta(f)) \leqslant \sigma_s(f,\varepsilon).$$

Proof. For a given $\varepsilon > 0$, we find a semialgebraic function g, such that $|f-g|_{C^1} \leqslant \varepsilon$. By definition of $|f-g|_{C^1}$ in Section 2 above, it follows that $|f(x)-g(x)| \leqslant \varepsilon$ for any $x \in B_r^n$, and $\|\nabla f(x) - \nabla(g)\| \leqslant \varepsilon/r$ for any $x \in B_r^n$, where g is smooth. Hence $\Sigma(f)$ is contained in the set of ε/r-critical points $\Sigma(g,\varepsilon/r)$ of g. In turn, $\Delta(f)$ is contained in an ε-neighborhood $\Delta_\varepsilon(g,\varepsilon/r)$. Therefore

$$M(\varepsilon, \Delta(f)) \leqslant M(\varepsilon, \Delta_\varepsilon(g,\varepsilon/r)) \leqslant C(g),$$

since the number of ε-intervals covering an ε-neighborhood of a bounded set in $\mathbb{R}$, is at most three times the number of ε-intervals covering the set itself.

Now taking infimum over all the semialgebraic g with $|f-g|_{C^1} \leqslant \varepsilon$, we get

$$M(\varepsilon, \Delta(f)) \leqslant \inf_g C(g) = \sigma_s(f,\varepsilon).$$

This completes the proof. □

Let us stress that the semialgebraic complexity $\sigma_s(f,\varepsilon)$ is a "correct" property of functions not only for the Sard Theorem itself, but (with minor modifications) for many related properties, investigated in [Yom1,Yom2,Yom3,Yom4,Yom5,Yom6,Yom7,Yom8,Yom17,Yom18,Yom-Com]: transversality results, average number of connected components of the fiber, etc. Also the "computational complexity" of most of the natural mathematical operations with f is bounded in terms of $\sigma_s(f,\varepsilon)$. In particular, this concerns the complexity of solving equations $f = \text{const}$ with a prescribed accuracy. This problem has been shortly discussed in [Yom18]. The bounds here can be obtained by a combination of the method of [Shu-Sma1,Shu-Sma2] (see also the continuing series of papers by the same authors) and of the above techniques. An implementation of the algorithm for solving nonlinear systems of equations, motivated by the semialgebraic complexity approach, has been constructed and preliminary tested [Eli-Yom4].

However, a very important exclusion in the applicability list of the semialgebraic complexity is given by some problems, arising in Dynamics. In particular, this concerns the results of [Yom13,Yom14,Yom15,Yom16]. To bound the complexity of the iteration of a mapping $f: M \to M$ (and, in particular, its topological entropy, volume growth, etc.), it is not enough to assume that $\sigma(f, \varepsilon)$ is small. The problem is that a piecewise-smooth structure of f, to which $\sigma(f, \varepsilon)$ is essentially insensible, in iterations can lead to an exponential growth of the number of smooth pieces, and thus to the blow up of the complexity. Consequently, we consider the following problem as an important one for understanding the nature of various complexity notions:

Is it possible to replace the C^k or analyticity assumptions in "dynamical" complexity bounds (like those obtained in [Yom13,Yom14,Yom15,Yom16]*) by weaker "complexity"-type assumptions?*

In the rest of this paper we present our main examples of functions, whose semialgebraic complexity is better (sometimes much better) than their regularity prescribes. These are maxima of smooth families (Section 4), compositions of smooth functions (Section 5), lacunary series (or Bernstein's quasianalytic classes)—Section 6, "Fewnomial series"—Section 7. Functions on infinite-dimensional spaces are considered in Section 8. This includes functions on ℓ^2, functions on $C^k[0,1]$ and certain control problems.

We believe that a richness of mathematical structures involved in these examples, and their potential applicability to many important problems in analysis justifies further investigation of semialgebraic complexity.

4. Maxima of smooth families

Let $h: B_r^n \times B_s^m \to \mathbb{R}$ be a continuous function. h can be considered as a family of functions $h_y: B_r^n \to \mathbb{R}$, $y \in B_s^m$. The (pointwise) maximum function $\max h$ of this family is defined as

$$\max h(x) = \max_{y \in B_s^m} h(x, y), \quad x \in B_r^n.$$

Functions of this form arise naturally in many problems of calculus of variation, optimization, control, etc. If we assume h to be C^k or analytic, it does not imply $\max h$ to be even once differentiable. (Notice, however, that for h a polynomial $\max h$ is semialgebraic and for h analytic $\max h$ is subanalytic. Lipschitz constant is preserved by taking maxima.)

Understanding the analytic nature of maxima functions is an important and mostly open problem. Besides its theoretical aspects, it presents a challenge in numerical applications: lack of smoothness of maxima functions prevents using many standard algorithms and packages. See, for example, [Roc,Sha-Yom,Yom6], where some partial results and references can be found. An important property of maxima functions is the fact that their singularities allow for a quite comprehensive description (see, for example, [Bry,Mat]). This property can be used in a numerical treatment of maxima functions, together with their low semialgebraic complexity, in the framework of our general approach.

The following result shows that while the regularity of h is usually completely lost in $\max h$, the complexity is preserved. To state it, we have to change a little bit our definition of semialgebraic complexity. Actually, C^1-assumption (as well as using the C^1-norm) in Definition 2.2 above have been made only to simplify a presentation and to avoid a necessity to define critical points and values for Lipschitzian functions.

Definition 4.1. Let $f : B_r^n \to \mathbb{R}$ be a continuous function. For any $\varepsilon > 0$, $\tilde{\sigma}_s(f, \varepsilon)$ is defined by

$$\tilde{\sigma}_s(f, \varepsilon) = \inf C(g),$$

where infimum is taken over all semialgebraic $g : B_r^n \to \mathbb{R}$ with $\max_{x \in B_r^n} |f(x) - g(x)| \leqslant \varepsilon$.

For any $f \in C^1$ we obviously have $\tilde{\sigma}_s(f, \varepsilon) \leqslant \sigma_s(f, \varepsilon)$, and usually the asymptotic behavior of $\tilde{\sigma}$ is the same as of σ, with $k - 1$ replaced by k.

Theorem 4.2. *Let $h : B_r^n \times B_s^m \to \mathbb{R}$ be a C^k-family and let $\max h$ be its maximum-function. Then*

$$\tilde{\sigma}_s(\max h, \varepsilon) \leqslant K \left(\frac{1}{\varepsilon}\right)^{\frac{n+m}{k}}.$$

Proof. Let $\varepsilon > 0$ be given. Find a polynomial $g : \mathbb{R}^{n+m} \to \mathbb{R}$, ε-approximating h on $B_r^n \times B_s^m$. By Jackson's theorem (see [Achi,Lor1,Lor2]), one can pick such a g with $\deg g \leqslant K_1 \left(\frac{1}{\varepsilon}\right)^{1/k}$. Let $\max g(x) = \max_{y \in B_s^m} g(x, y)$ be the maximum function of the family g.

For any $x \in B_r^n$, $|\max h(x) - \max g(x)| \leqslant \varepsilon$. Indeed, $\max h(x) = \max_y h(x, y) = h(x, \bar{y})$ for a certain $\bar{y} \in B_s^m$. Now, $\max g(x) = \max_y g(x, y) \geqslant g(x, \bar{y})$. But $|h(x, \bar{y}) - g(x, \bar{y})| \leqslant \varepsilon$, and hence

$$\max h(x) \leqslant \max g(x) + \varepsilon.$$

Changing the places of h and g, we get in the same way

$$\max h(x) \geqslant \max g(x) - \varepsilon,$$

or, finally, $|\max h(x) - \max g(x)| \leqslant \varepsilon$.

$\max g$ is a semialgebraic function (define $\max g$ via projections of the graph of g and apply the Tarskii–Zeidenberg theorem—see e.g. [Boc-Cos-Roy]). To continue the proof we need the following result:

Lemma 4.3. *The "Sard complexity measure" $C_s(\max g)$ satisfies*

$$C_s(\max g) \leqslant K_2(\deg g)^{n+m}.$$

Proof. Let $x \in B_r^n$ be a point where $\max g$ is analytic. It is easy to see that there is $y \in B_s^m$ such that $\max g(x) = g(x,y)$, $\nabla \max g(x) = \nabla_x g(x,y)$ and $\nabla_y g(x,y) = 0$. Hence, for x an ε-critical point of $\max g$, (x,y) is an ε-critical point of g. Respectively, $\max g(x) = g(x,y)$ is an ε-critical value of g. The result now follows from Theorem 3.2 stated above (or rather from a version of this theorem, where the ball is replaced by the product of two balls). This proves the lemma.

To complete the proof of Theorem 4.2 it remains to use the fact that by the construction above, $\deg g \leqslant K_1\left(\frac{1}{\varepsilon}\right)^{1/k}$. Hence

$$C(g) \leqslant K_2(\deg g)^{n+m} \leqslant K_3\left(\frac{1}{\varepsilon}\right)^{\frac{m+n}{k}}. \qquad \square$$

Remark 1. The result of Theorem 4.2 can be used exactly as in Theorem 3.6 above, in order to estimate the ε-entropy of critical values attained at *smooth* critical point of $\max h$. However, the maxima of smooth families may have singularities, and it is important to control also "singular" critical points and values.

An important fact is that critical (and near-critical) points and values can be defined for general Lipschitzian function, using the notion of Clarke's generalized differential (see [Cla1,Cla2,Yom9,Yom10,Yom11,Yom12]). A corresponding version of Theorem 4.2 above remains true. However, its proof becomes much more tricky, and we do not give details here. Different versions of quantitative Sard theorem for maximum functions and for continuous selections can be found also in [Roh1,Roh2,Roh3].

Remark 2. Essentially, the same method as above proves a much more general result than Theorem 4.2: *the semialgebraic complexity of the maximum function* $\max h$ *can be bounded through the semialgebraic complexity of the family* h. However, technically this proof is much more involved, and we plan to present it separately.

Remark 3. As far as the generalized derivatives of the maximum functions are concerned, we can consider two versions of the general conjecture of Section 2. The first version corresponds to the complexity bound of Theorem 4.2: *The maximum function* $\max h$ *has a* $\frac{nk}{n+m}$*th Peano differential almost everywhere*. The second version is more optimistic. It is motivated by some special examples of the maxima of smooth families: *The maximum function* $\max h$ *of a* C^k*-family* h *has a* k*th Peano differential almost everywhere*, *independently of the dimension of the family*.

5. Semialgebraic complexity of functions representable by compositions

A possibility to represent a given function of several variables as a chain (or a composition) of functions of a smaller number of variables, apparently presents a basic restriction on the nature of this function. Still it turns out to be a difficult problem to find analytic or topological obstructions to such a representability. We refer the reader to Vitushkin's paper [Vit3] for a survey of the long history of this problem and (to some extent) of its current status.

The degree of differentiability of a composition is equal to the minimum of the degrees of differentiability of the functions in the chain. In [Vit1] Vitushkin has shown that this condition is not sufficient for a composition representation (see also [Vit2,Kol-Tih,Tih]). By certain "massiveness" arguments he proved that under some conditions on the composition scheme, most of the functions of the required smoothness are not composition representable.

Semialgebraic complexity turns out to be quite sensitive to a composition representability of functions. In the examples considered below (and in many other cases) a possibility to represent a function via a certain composition scheme implies a specific upper bound on its semialgebraic complexity. In particular, this gives a tool to produce *explicit examples* of nonrepresentable functions.

Assuming that a function is representable according to a certain composition scheme, we can try to estimate its semialgebraic complexity as follows: (a) We approximate the components of the representation by semialgebraic functions (in particular by piecewise polynomials). (b) Composing these semialgebraic functions, we get a semialgebraic approximation of f, whose complexity can be bounded, using the specifics of the composition scheme.

In some cases the resulting bound on $\sigma_s(f, \varepsilon)$ is better than the regularity of f prescribes. As it happens, two important conclusions can be derived:

First, that the composition scheme considered provides nontrivial restrictions on the complexity of representable functions. This fact can be used both theoretically and numerically (suggesting, for example, that a numerical approximation of functions of such a form should explicitly take into account their composition structure).

Second, one can easily produce explicit examples of functions, not representable according to this composition scheme: it is enough to take f of a required regularity, with $\sigma_s(f, \varepsilon)$ bigger than the composition bound prescribes.

Notice that "massiveness" methods, mentioned above, prove existence of nonrepresentable functions, without producing any specific example.

Below we bound $\sigma_s(f, \varepsilon)$ for two rather special types of composition representations. A natural question is: *for what other types of compositions* $\sigma_s(f, \varepsilon)$ *is better than prescribed by the regularity*?

5.1. Generalized Kolmogorov's representation

One of the most striking results in the composition problem was the following theorem of Kolmogorov [Kol]:

Any continuous function of n variables can be represented as

$$f(x_1, x_2, \dots, x_n) = \sum_{q=1}^{2n+1} g_q\left(\sum_{p=1}^{n} \varphi_{p,q}(x_p)\right), \tag{5.1}$$

with g_q—continuous functions of one variable, depending on f, and $\varphi_{p,q}$—universal continuous monotonic functions of one variable.

In a composition scheme (5.1) one can now assume a higher regularity of the components, and to estimate the complexity of the representable functions. We shall slightly generalize (5.1) as a composition scheme, but impose a very strong regularity conditions, in order to simplify computations. Consider a composition scheme for functions $f : B_1^n \to \mathbb{R}$:

$$f(x_1, x_2, \dots, x_n) = \sum_{i=1}^{m} g_i(P_i(x_1, \dots, x_n)), \tag{5.2}$$

where P_i are polynomials of degree d in $x_1, \dots, x_n$ bounded by 1 on B_1^n, and g_i are C^k functions of one variable, bounded by 1 together with all their derivatives up to order k on the interval $[-1, 1]$. Any f, representable by (5.2), is a C^k function, and unless some specific relations between the components are satisfied, it is not C^{k+1}.

Theorem 5.1. *For any f, representable by* (5.2),

$$\sigma_s(f, \varepsilon) \leqslant K\left(\frac{1}{\varepsilon}\right)^{m/k},$$

where K depends only on m, k and d.

Proof. For a given $\varepsilon > 0$, we $\frac{\varepsilon}{m}$-approximate g_i by their Taylor polynomials Q_i^j of degree k on the subintervals Δ_j of $[-1, 1]$ of the length $K_1\varepsilon^{1/k}$, $j = 1, 2, \dots, ([2/k_1] + 1)\left(\frac{1}{\varepsilon}\right)^{1/k}$.

Polynomials $Q_i^j(P_i)$ (each of degree kd) $\frac{\varepsilon}{m}$-approximate $g_i(P_i)$ on semialgebraic subsets Ω_i^j of B_1^n, $\Omega_i^j = P_i^{-1}(\Delta_j)$. Finally, for any $j_1, \dots, j_m$, the polynomial $\sum_{i=1}^m Q_i^{j_i}(P_i)$ of degree kd, ε-approximates f on a semialgebraic subset $\Omega^{j_1,\dots j_m} = \Omega_1^{j_1} \cap \dots \cap \Omega_m^{j_m}$ of B_1^n. The complexity of the resulting piecewise-polynomial approximating function g, $C(g)$, is bounded by $(kd)^n \cdot K_3\left(\frac{1}{\varepsilon}\right)^{m/k}$, according to Theorem 3.5 above, since the number of semialgebraic pieces $\Omega^{j_1,\dots,j_m}$ is $K_2\left(\frac{1}{\varepsilon}\right)^{m/k}$, each piece is defined by $2m$ polynomial inequalities of degree d, and the degree of the approximating polynomial on each piece is kd. This proves the theorem. $\square$

Applying Theorem 3.6 we get the following result:

Corollary 5.2. *For any f, representable by* (5.2),

$$M(\varepsilon, \Delta(f)) \leqslant K\left(\frac{1}{\varepsilon}\right)^{m/k}.$$

The exponent of $\frac{1}{\varepsilon}$ in the rate of the growth of the ε-entropy is called the entropy dimension $\dim_e$ of the set. We get that the entropy dimension of the set of critical values of f, $\dim_e \Delta(f)$, does not exceed $\frac{m}{k}$ for any f, representable by (5.2).

Corollary 5.3. *If in* (5.2) $m < n$, *then there exist C^k-functions f on B_1^n, not representable according to* (5.2). *In fact, any f with* $\dim_e \Delta(f) > \frac{m}{k}$, *is not representable by* (5.2).

Examples of smooth functions with a large set of critical values (i.e. with $\dim_e \Delta = \frac{n}{k}$) are produced by the Whitney construction [Whi1]. (A natural general question here is *whether for the "most" of C^k-functions f* $\dim_e \Delta(f)$ *is* $\frac{n}{k}$?)

Another question that arises is: *What other obstructions* (*besides smoothness and complexity*) *exist to a representability of functions by* (5.2)?

Remark. If we approximate g_i above by global polynomials, we get for $\varepsilon > 0$ given, the degree of ε-approximating polynomial $\tilde{g}_i$ of order $\left(\frac{1}{\varepsilon}\right)^{1/k}$, and the degree of the compositions $\tilde{g}_i(P_i)$ and of their sum equal to $d\left(\frac{1}{\varepsilon}\right)^{1/k}$. As one can expect, we get no improvement to the bound, provided by the C^k-smoothness of f: indeed, we remain in the framework of global polynomial approximations, whose rates are essentially equivalent to the regularity.

5.2. Direct compositions

For $f : B^{n_1} \to B^{n_{r+1}}$ we call "direct composition" a representation according to the following scheme:

$$f = f_r \circ f_{r-1} \circ \cdots \circ f_1 \tag{5.3}$$

with $f_i : B^{n_i} \to B^{n_{i+1}}$. For each $i = 1, \ldots, r$, $f_i \in C^{k_i}$, with fixed $k_i \geqslant 2$. Here $n_1 = n$, and we always assume that $n_1 \geqslant n_2 \geqslant \cdots \geqslant n_r \geqslant n_{r+1}$.

We can assume additionally that $k_1 > k_2 > \cdots > k_r$. Indeed, consider a composition of two mappings $\varphi = f_{j+1} \circ f_j$. The following lemma shows that if $k_{j+1} \geqslant k_j$, then in such a form can be represented exactly all the C^{k_j}-mappings from $\mathbb{R}^{n_j}$ to $\mathbb{R}^{n_{j+2}}$. Therefore we can replace the couple $(j, j+1)$ in the direct composition scheme (5.3) by one C^{k_j}-mapping.

Lemma 5.4. *For $r = 2$ and $k_2 \geqslant k_1$, φ can be represented according to a scheme* (5.3) *if and only if it belongs to C^{k_1}.*

Proof. In one direction this is obvious: a composition of the above form is always in C^{k_1}. Conversely, if $\varphi \in C^{k_1}$, it can be represented as $\varphi = \varphi_2 \circ \varphi_1$, with $\varphi_1(x_1, \dots, x_{n_1}) = (\varphi(x_1, \dots, x_{n_1}), 0, \dots, 0)$ and φ_2 the projection of $\mathbb{R}^{n_2}$ on the first n_3 coordinates.

However, under the assumption $k_1 > k_2 > \dots > k_r$, for $f \in C^{k_r}$ its representability by (5.3) may be a nontrivial restriction. This depends on the dimensions n_j and on the degrees of differentiability k_j. Some restrictions here can be obtained by "massiveness arguments" along the lines of Kolmogorov's and Vitushkin's approaches [Kol-Tih,Vit1,Vit2,Vit3,Tih]. We shall bound instead the semialgebraic complexity of f, and as a result obtain explicit obstructions to a direct composition representability.

Since in this paper we have introduced the notion of a semialgebraic complexity only for functions (and not for mappings into higher-dimensional spaces) we have to assume below that in a composition scheme (5.3), $f : B_1^n \to \mathbb{R}$, and hence $n_{r+1} = 1$. $\square$

Theorem 5.5. *For any f, represented by* (5.3), $\sigma_s(f, \varepsilon) \leqslant C\left(\frac{1}{\varepsilon}\right)^{\delta(f)}$, *where*

$$\delta(f) = \frac{n_1 - n_2}{k_1 - 1} + \frac{n_2 - n_3}{k_2 - 1} + \dots + \frac{n_{r-1} - n_r}{k_{r-1} - 1} + \frac{n_r}{k_r - 1}.$$

Proof. We give only a sketch. We plan to present a detail proof (for much more general compositions schemes) separately. See also [Yom8], where a similar technique is used for a direct estimation of the size of critical values of compositions, without computing their semialgebraic complexity.

As usual, for a given $\varepsilon > 0$, we ε-approximate each f_j by a piecewise-polynomial mapping, formed by k_j-order Taylor polynomials of f_j on subcubes of $B_1^{n_j}$ of the size $K_1 \varepsilon^{1/k_j - 1}$. Composition g of these mappings provides a piecewise-polynomial $K_2\varepsilon$-approximation of f. To estimate the number of the elements in the partition of B_1^n, on which this composition is piecewise-polynomial, we notice that the image under (the approximation of) f_j of any cube in $B_1^{n_j}$ of the size $K_1 \varepsilon^{1/k_j - 1}$ is contained in a certain cube of the size $K_3 \varepsilon^{1/k_j - 1}$ in $B_1^{n_{j+1}}$. Hence this image can intersect at most $K_4 \varepsilon^{n_{j+1}\left(\frac{1}{k_j - 1} - \frac{1}{k_{j+1} - 1}\right)}$ cubes of the partition of $B_1^{n_{j+1}}$. Hence the total number of the composition chains on different partition boxes is bounded by K_5 times $1/\varepsilon$ to the power

$$\frac{n_1}{k_1 - 1} - \frac{n_2}{k_1 - 1} + \frac{n_2}{k_2 - 1} - \frac{n_3}{k_2 - 1} + \frac{n_3}{k_3 - 1} - \dots + \frac{n_{r-1}}{k_{r-1} - 1} - \frac{n_r}{k_{r-1} - 1} + \frac{n_r}{k_r - 1}$$
$$= \frac{n_1 - n_2}{k_1 - 1} + \frac{n_2 - n_3}{k_2 - 1} + \dots + \frac{n_{r-1} - n_r}{k_{r-1} - 1} + \frac{n_r}{k_r - 1} = \delta.$$

The degrees of the polynomials on each piece are fixed (and equal to $k_1 \cdot k_2 \cdots k_r$) and hence the complexity $C(g)$ is of order $(1/\varepsilon)^{\delta}$, according to Theorem 3.5 above. This completes the proof. $\square$

Corollary 5.6. *If in* (5.3) $\delta < \frac{n_1}{k_r}$, *then there exist* C^{k_r}*-functions, not representable according to* (5.3). *In fact, any* f *with* $\dim_e \Delta(f) > \delta$ *is not representable.*

The proof is exactly the same as for the Corollaries 5.2 and 5.3.

Now one can find many sequences $n_1 \geqslant n_2 \geqslant \cdots \geqslant n_r$ and $k_1 > k_2 > \cdots > k_r$, with $\delta < \frac{n_1}{k_r}$. For example, for $n = n_1 = 10$, $n_2 = 5$, $k_1 = 52$, and $k_2 = 8$, we have $\delta = \frac{5}{51} + \frac{5}{7} = \frac{295}{364} < \frac{5}{6}$. Hence the C^9-Whitney function $h : B_1^{10} \to \mathbb{R}$, which has $\Delta(h) = [0, 1]$, cannot be represented by (5.3). In the same way, C^{11}-functions f with $\dim_e \Delta(f) = \frac{10}{11} > \delta$ cannot be represented by (5.3).

6. Lacunary series

The approach of Bernstein to real analytic and quasianalytic functions starts with their polynomial approximations [Ber1,Ber2,Ber3]. In particular, it is shown in [Ber2] that *a real function* $f(x)$, *defined on the closed interval* $[a, b]$ *is real analytic on it if and only if there are constants* $M > 0$ *and* $\rho < 1$ *so that for any natural* d, $E_{\text{poly}}(f, d) \leqslant M\rho^d$ *on* $[a, b]$. (See Section 2 above for a definition of $E_{\text{poly}}(f, d)$.)

Bernstein defines (P)-quasianalytic function as follows [Ber3]:

Definition 6.1. $f(x)$ is quasianalytic (P) on $[a, b]$, if

$$E_{\text{poly}}(f, d) \leqslant M\rho^d \quad \text{on} \quad [a, b] \tag{6.1}$$

for an infinite sequence of degrees d. The sequence $d_1, d_2, \ldots, d_i, \ldots$ of the degrees, for which (6.1) is satisfied, is called a (P)-sequence of f.

(P)-quasianalytic functions may be analytic in the usual sense (this always happens if the (P)-sequence of f is not *lacunary*, i.e. if the ratio d_{i+1}/d_i is uniformly bounded). This function is C^∞ if $(\log d_{i+1})/d_i \to 0$, and f may have only a finite number of derivatives or not to be differentiable at all for $(\log d_{i+1})/d_i \geqslant \text{const} > 0$. However, in any case (P)-quasianalytic functions have a property of a "quasianalytic continuation": they are uniquely defined by their values on any subinterval of $[a, b]$ (see [Ber3]). Comparison of (P)-quasianalytic functions with Danjoi–Carleman quasianalytic functions can be also found in [Ber3] and other papers by Bernstein.

From the point of view of a semialgebraic complexity, the restriction of the approximating sequence of polynomials (semialgebraic functions) to a lacunary subsequence of the degrees, does not change much. Most of the constructions of this paper can be modified accordingly. It is not natural in this setting to restrict ourselves to the "analytic approximation rate" $E_{\text{poly}}(f, d) \leqslant M\rho^d$ only: nontrivial complexity bounds manifest themselves also for a "C^k-approximation rate" $E_{\text{poly}}(f, d) \leqslant (1/d)^k$.

So assume that a function $f : B_r^n \to \mathbb{R}$ can be fastly approximated by polynomials of degree d only for d in a certain subsequence $d_1, d_2, \ldots, d_i$, of natural numbers.

Consider a sequence $\varepsilon_i = E_{\text{poly}}(f, d_i)$, i.e. ε_i is the error of the best approximation of f in a C^1-norm by polynomials of degree d_i. Then it is natural to consider a (polynomial) complexity $\sigma_{\text{poly}}(f, \varepsilon)$ only for the values $\varepsilon = \varepsilon_i$. Exactly in the same way we can consider a semialgebraic complexity $\sigma_{\text{s}}(f, \varepsilon)$, taking $\varepsilon = \varepsilon_i$. This motivates the following definition:

Definition 6.2. Let a sequence $E = (\varepsilon_1 > \varepsilon_2 > \cdots > \varepsilon_i > \cdots)$ of positive numbers, tending to zero, be given. We call an E-complexity (semialgebraic or polynomial, respectively) the restriction of $\sigma_{\text{s}}(f, \varepsilon)$ (of $\sigma_{\text{poly}}(f, \varepsilon)$, respectively) to the sequence E.

Now it is a direct consequence of Theorem 3.6 above, that *for* $\varepsilon \in E$, ε*-entropy of* $\Delta(f)$ *is bounded from above by the* E*-complexity of* f.

We consider as a very interesting problem the investigation of such sets Δ, whose ε-entropy behaves "well" (in a "C^k" or "analytic" way) only for a lacunary sequence of ε_i. In particular, *is it possible to bound the entropy dimension of such sets*? *What is their characterization in terms of the "*β*-spread"* (*as defined in* [Yom4])? *How to extend standard examples and constructions in Fractal Geometry* (*see*, *for example*, [Tri]) *to such "lacunary fractals"*?

However, to estimate the usual Lebesgue measure μ of Δ (as well as its Hausdorff measure and its Hausdorff dimension) the "lacunary" information above is sufficient.

Theorem 6.3. *If for a certain* E *as above* E*-complexity of* f *is bounded by* $K(1/\varepsilon)^{\delta}$, $\delta < 1$, *then* $\mu(\Delta(f)) = 0$.

Proof. Clearly
$\mu(\Delta(f)) \leqslant \inf_{\varepsilon > 0} \varepsilon \cdot M(\varepsilon, \Delta(f)) \leqslant \inf_{\varepsilon \in E} \varepsilon M(\varepsilon, \Delta(f)) \leqslant \inf_{\varepsilon \in E} \varepsilon \cdot K(1/\varepsilon)^{\delta} = 0.$ $\square$

Remark. A similar computation allows one to bound the Hausdorff measure of $\Delta(f)$ in any dimension. In particular, it shows that the Hausdorff dimension of $\Delta(f)$ is at most δ.

Question. *Does the "lacunary complexity" bound the geometry of the level sets of* f?

The usual semialgebraic complexity does, according to [Yom3] and Section 9.1 below. This concerns the average Betti numbers, average curvature bounds for the level sets, etc.

Example. Consider the function $f : B_1^n \to \mathbb{R}$ of the form

$$f(x) = \sum_{i \in D} a_i p_i(x), \tag{6.2}$$

where D is a certain sequence $d_1, d_2, \dots, d_j, \dots$ of the degrees, $P_i : \mathbb{R}^n \to \mathbb{R}$ are polynomials of degree i. We'll assume that the series

$$\sum_{i \in D} i^2 |a_i| \max_{x \in B_1^n} |p_i(x)|$$

converges. Since by the Markov inequality [Achi], $|\nabla p_i|$ on B_1^n is bounded by $K \cdot i^2 \max_{x \in B_1^n} |p_i(x)|$, this assumption implies that $f \in C^1$.

Moreover, for any d_j, the polynomial $\sum_{i \in D, i \leqslant d_j} a_i p_i(x)$ has degree d_j, and its deviation in C^1-norm from f does not exceed $\delta_j = \sum_{i \in D, i > d_j} (i^2|a_i| + 1) \max_{x \in B_1^n} |p_i(x)|$. Thus we have the following bound:

Proposition 6.4. *For f defined by* (6.2) *and for $\varepsilon = \delta_j$, $j = 1, 2, \dots$*

$$\sigma_s(f, \delta_j) \leqslant \sigma_p(f, \delta_j) \leqslant K d_j^n.$$

Let us assume now that $p_j(x)$ are normalized: $\max_{x \in B_1^n} |p_j(x)| = 1$. To simplify computations, we assume also that the sequence of the coefficients a_i is fixed (say, $a_i = 1/i^4$), and analyze only the dependence of the E-complexity of f on the sequence d_j. We have $\delta_j = \sum_{i \in D, i > d_j} 1/i^2 \leqslant K_1(1/d_{j+1})$. Consequently, $\sigma_p(f, \delta_j) \leqslant d_j^n = (1/\delta_j)^{\alpha_j}$, with $\alpha_j \sim (n \log d_j / \log d_{j+1})$.

Corollary 6.5. *If the sequence d_j satisfies: for any j, $(\log d_{j+1}/\log d_j) \geqslant \beta > n$, then the set of critical values of f has measure* 0.

The condition $(\log d_{j+1}/\log d_j) \geqslant \beta > n$ is satisfied, for example, if $d_{j+1} = d_j^\beta$, $\beta > n$. Notice that in this example f is not (P)-quasianalytic, since its approximation rate on the lacunary sequence of the degrees d_j is of a "$C^{n/\beta}$" type. If we assume that a_i decrease faster, the same result will be obtained for a "less lacunary" sequence of d_j.

It is interesting to compare the influence of the lacunarity of the sequence of d_j on the regularity in the classical Bernstein's setting and on the semialgebraic complexity in the example above. In Bernstein's definition stronger lacunarity of the sequence d_j reduces the regularity. On the contrary, in our example stronger lacunarity implies lower complexity (at least as the size of the critical values is concerned). This apparent contradiction is explained by the fact that while in Bernstein's the rate of approximation is fixed, in our example this rate grows with the lacunarity.

As far as the Sard theorem is concerned, the fact that in the above examples polynomial sums are used, is not essential. We can replace everywhere polynomials of degree d by C^1 semialgebraic functions of the "Sard complexity" d^n. Corollary 6.5 and its proof do not change.

7. Fewnomials

In this section we give an additional illustration of the fact that the semialgebraic complexity is sensitive to rather subtle analytic properties of functions, which may not be detected by the usual approximation methods. It concerns the rate of approximation by the so-called "fewnomials".

It was conjectured by Arnold and later proved by Khovanskii [Kho] that the number of real solutions of a polynomial system (in a positive octant, properly counted) is bounded by the number of the nonzero monomials in the system (and not by the degree!) In most of the results in real semialgebraic geometry, relevant to our approach, the geometric bounds on semialgebraic sets (their volumes, ε-entropy, variations) are obtained via a reduction to the counting of isolated real solutions of a certain polynomial system. Consequently, one can hope to build a version of the theory, where the degrees are replaced by the number of nonzero monomials.

Although in most of analytic applications it is difficult to expect that the approximating polynomials will have a few nonzero terms, we believe that in some cases the "fewnomial" version of our approach can be very useful. Constructing this "fewnomial" version is not completely straightforward, and we consider it as an important open problem. Of course, other complexity measures of polynomials (like "additive complexity") can be used, and it would be interesting to compare the resulting versions of the semialgebraic complexity.

The following example shows that the discussed extension may be useful even in the simplest situations of functions of one variable. For polynomials of one variable the "fewnomials" theory is represented by the classical Descartes lemma:

The number of positive roots of a real polynomial does not exceed the number of nonzero terms in it.

For a real polynomial $p(x) = \sum a_i x^i$ of one variable let us denote $q(p)$ the number of the nonzero coefficients a_i. Thus the number of the positive zeroes of f is bounded by $q(p)$.

Let us prove here a simple "fewnomial" version of the Quantitative Sard theorem.

Lemma 7.1. *Let $p(x)$ be a real polynomial, restricted to the interval $[0, 2]$. Then for the set of γ-critical values $\Delta(p, \gamma)$ of f on $[0, 2]$, the following inequality is satisfied:*

$$M(\varepsilon, \Delta(p, \gamma)) \leqslant 2q(p)\left(1 + \frac{\gamma}{\varepsilon}\right).$$

In particular, for the ε-critical values and for the usual critical values

$$M(\varepsilon, \Delta(p)) \leqslant M(\varepsilon, \Delta(p, \varepsilon)) \leqslant 4q(p).$$

Proof. The set $\sum(p, \gamma)$ of the γ-critical points of p is defined by $|p'| \leqslant \gamma$. But p' also contains at most $q(p)$ nonzero terms, and hence by the Descartes lemma, each of the equations $p' = \gamma$ and $p' = -\gamma$ can have at most $q(p)$ solutions on $[0, 2]$. Hence $\sum(p, \gamma)$ consists of at most $q(p)$ intervals inside $[0, 2]$. Since $|p'| \leqslant \gamma$ on each of these intervals, their images under p are the intervals of the length at most 2γ.

Hence $\Delta(p,\gamma)$ is a union of $q(p)$ intervals of the length at most 2γ, and it can be covered by not more than $q(p)\left(1+\frac{\gamma}{\varepsilon}\right)$ 2ε-intervals. □

Corollary 7.2. *For $p(x)$ as above the "Sard complexity" $C(p)$ is bounded by $4q(p)$.*

Let us now give the "fewnomial" version of Theorem 3.6. For a real C^1-function f on $[0,2]$ define a "fewnomial" complexity $\sigma_{\text{few}}(f,\varepsilon)$ as

$$\inf_{|f-p|_{C^1}\leqslant\varepsilon} 4q(p),$$

where the infimum is taken over all the real polynomials p. Now applying Corollary 7.2 instead of Theorem 3.2, we get a "fewnomial" version of Theorem 3.6:

Theorem 7.3. *For f as above and for any $\varepsilon>0$,*

$$M(\varepsilon,\Delta(f))\leqslant\sigma_{\text{s}}(f,\varepsilon)\leqslant\sigma_{\text{few}}(f,\varepsilon).$$

Examples of f with a prescribed fewnomial complexity can be constructed in the form

$$f(x)=\sum_{i=0}^{\infty} a_i p_i(x),$$

where p_i is a fixed sequence of polynomials with, say, $q(p_i)=i$. Then

$$q\left(\sum_{i=0}^{n} a_i p_i\right)\leqslant\frac{n(n+1)}{2},$$

and chosing an appropriate rate of converging of the coefficients a_i to zero, we get a prescribed rate of a fewnomial approximation. In particular, if we assume that the maximum of the polynomials p_i and of their derivatives on $[0,2]$ are bounded by 1, and use as the approximating fewnomials the partial sums

$$P_n(x)=\sum_{i=0}^{n} a_i p_i(x),$$

we get

$$|f-P_n|_{C^1}\leqslant\varepsilon_n=2\sum_{i=n+1}^{\infty}|a_i|.$$

To simplify notations, let us assume that $a_i=\left(\frac{1}{i}\right)^{k+1}$. Then ε_n is of order $\left(\frac{1}{n}\right)^k$, or $n\sim\left(\frac{1}{\varepsilon_n}\right)^{\frac{1}{k}}$. Since $q(P_n)\leqslant\frac{n(n+1)}{2}$, we finally get:

Theorem 7.4. *For f as above*

$$M(\varepsilon, \Delta(f)) \leqslant \sigma_s(f, \varepsilon) \leqslant \sigma_{\text{few}}(f, \varepsilon) \leqslant K\left(\frac{1}{\varepsilon}\right)^{\frac{2}{k}}.$$

In particular, $\dim_e(\Delta(f)) \leqslant \frac{2}{k}$.

It is interesting to compare the result of Theorem 7.2 with the conclusion of the usual Quantitative Sard theorem. If the polynomials p_i in the sum above have degree i, then under the same assumptions f is $k+1$ times differentiable. Hence, $\dim_e(\Delta(f)) \leqslant \frac{1}{k+1}$, by Theorem 3.1. Theorem 7.4 states that replacing the degree of p_i by the number of the nonzero monomials is equivalent (as the critical values are concerned) to the reducing twice the differentiability of f. It would be interesting to describe *analytic properties of functions with a prescribed rate of the fewnomial approximation.*

Lacunary series, considered in Section 6 above, can be investigated also from the "fewnomial" point of view. Consider a function of one variable $f(x)$ on the interval $I = [0, 2]$, given by

$$f(x) = \sum_{i \in D} a_i p_i, \quad D = (d_1, d_2, \dots). \tag{7.1}$$

Here p_i satisfy the same assumptions as above (i.e. we assume that the maximum of the polynomials p_i and of their derivatives on $[0, 2]$ is bounded by 1 and that the number of the nonzero terms in p_i, $q(p_i)$, is at most i). Since we are interested in the dependence of the complexity of f on the set D of the indices, let us fix, as in Section 6, $a_i = 1/i^4$. As above, f is approximated in a C^1-norm on I by the polynomial $\varphi_j = \sum_{i \in D, i \leqslant d_j} a_i p_i$ with the accuracy at least

$$\delta_j = 2 \sum_{i \in D, i > d_j} |a_i| \sim \left(\frac{1}{d_{j+1}}\right)^3. \tag{7.2}$$

Now, φ_j contains at most $\frac{d_j(d_j+1)}{2}$ nonzero terms. Hence we get the following:

Proposition 7.5. *For f given by* (7.1) *and δ_j defined by* (7.2), $j = 1, 2, \dots,$

$$\sigma_{\text{few}}(f, \delta_j) \leqslant K d_j^2.$$

Comparing this result with Proposition 6.4 (where n is equal to 1) we see that also in the lacunary case replacing the degree with the number of the nonzero terms "reduces twice the smoothness".

However, we can assume that $q(p_i)$ grows with the index slower than i. For example, assume that $q(p_{d_j})$ is at most j. Proposition 7.5 takes the following form:

Proposition 7.6. *For f given by* (7.1) *with $q(p_{d_j}) \leqslant j$, and δ_j defined by* (7.2), $j = 1, 2, \dots,$

$$\sigma_{\text{few}}(f, \delta_j) \leqslant K j^2.$$

Let us now repeat the computations in the end of Section 6, in order to find the distributions of the degrees d_j with the fewnomial complexity of f, corresponding to a C^k-regularity, $k>0$ (at least from the point of view of the size of critical values).

Since $\delta_j \sim (1/d_{j+1})^3$, we get $d_{j+1} \sim (1/\delta_j)^{1/3}$. Denote by ψ the inverse function to d_j: $j = \psi(d_j)$. Thus

$$\sigma_{\text{few}}(f, \delta_j) \leqslant K(j+1)^2 \sim \psi\left(\frac{1}{\delta_j}\right)^{2/3}.$$

We see that the fewnomial complexity $\sigma_{\text{few}}(f, \delta_j)$ behaves as $(1/\delta_j)^{\alpha}$ for ψ having a form $\psi(z) = z^{\beta}$, $\alpha = \frac{2}{3}\beta$, and we have $\alpha < 1$ for $\beta < 3/2$. We get the following result:

Corollary 7.7. *If the degrees d_j grow with j faster than j^{γ}, $\gamma > \frac{2}{3}$, then $m(\Delta(f)) = 0$.*

It would be very interesting to compare various complexities, described above. In particular, the following problem naturally arises:

Describe in analytic terms (or at least give nontrivial analytic restrictions on) the classes of functions of a given semialgebraic (fewnomial, additive,...) complexity. Give examples of functions of a high fewnomial complexity, but with a low semialgebraic one. Are there natural classes of functions, for which all the complexities above behave in a similar way? What specific properties of regular functions (C^k, analytic) are enforced by a requirement of their low complexity?

8. Complexity of functions on infinite-dimensional spaces

The notion of semialgebraic complexity presented in this paper applies equally well to functions on infinite-dimensional spaces. Moreover, it is in this setting that the difference between the notions of complexity and regularity becomes apparent. This due to the fact that a "polynomial of an infinite number of variables" is not a "simple function" (unless it satisfies some additional restrictions). Thus just a regularity, which in the infinite-dimensional situation is still roughly equivalent to the rate of a global polynomial approximation, provides no information on the complexity. This is in a strong contrast to the finite-dimensional case—see Theorem 2.3 above.

Let us start with an example of a polynomial on ℓ^2, which does not satisfy the Sard theorem (this example belongs to Kupka [Kup]).

Let $\ell^2 = \{x = (x_1, \dots, x_i, \dots), \Sigma x_i^2 < \infty\}$ be the standard Hilbert space. We define a function

$$f : \ell^2 \to \mathbb{R} \quad \text{by} \quad f(x) = \sum_{i=1}^{\infty} \frac{1}{2^i} \varphi(ix_i), \tag{8.1}$$

where φ is the polynomial of one variable of degree 3, such that $\varphi(0) = \varphi'(0) = 0$, $\varphi(1) = 1$ and $\varphi'(1) = 0$. Thus φ has exactly two critical points 0 and 1 with the critical values 0 and 1, respectively. One can easily show that f is infinitely differentiable (in any reasonable definition of a differentiability on ℓ^2). In fact, f can be considered as an analytic function, or even as a polynomial of degree 3 on ℓ^2. Now $x = (x_1, \dots, x_i, \dots)$ is a critical point of f if and only if for each i, ix_i is a critical point of φ. Thus

$$\Sigma(f) = \left\{ \left(a_1, \frac{a_2}{2}, \dots, \frac{a_i}{i}, \dots \right), a_i = 0, 1 \right\} \quad \text{and}$$
$$\Delta(f) = \left\{ \sum_{i=1}^{\infty} \frac{1}{2^i} a_i, a_i = 0, 1 \right\} = [0, 1].$$

Hence the critical values of a C^∞-function f cover the interval: the Sard theorem is no more valid on ℓ^2.

There are some infinite-dimensional examples (usually assuming a certain "compactness" of the considered operators) where the Sard theorem is still valid (see [Sma]). However, in general one cannot expect that infinite smoothness or even analyticity is sufficient, as the above example shows. To our best knowledge no general criteria for a validity of the Sard theorem on infinite-dimensional spaces has been suggested. We've noticed above that a deep reason for this situation is that the validity of the Sard theorem for a certain function is a manifestation of its complexity, and not of its regularity. While in finite dimension these two notions roughly agree, in infinite-dimensional case they become almost independent (see [Zer]).

Notice that the function f of the Kupka example can be approximated by "simple" ones, namely by the polynomials, depending only on a finite number of variables: $\sum_{i=1}^{N} \frac{1}{2^i} \varphi(ix_i)$. Developing further this observation, we shall show that f violates the Sard theorem since the rate of its approximation by these "simple" polynomials is not high enough.

Kupka's example shows that in order to apply the approach of this paper to the infinite-dimensional situation, we have first to find a good class of "simple" approximating functions. This suggests the following generalization of our main Definition 2.2:

Let V be a Banach space (of finite or infinite dimension) and let $B \subset V$ be a closed set in V (mostly B will be the unit ball). We consider (Frechet) continuously differentiable functions on B and the norm $| \ |_{C^1}$ as defined in Section 2.

Now assume that some subclass Q of such functions is given, satisfying the following condition $(*)$: *Fix any $q \in Q$. Then for any $\varepsilon > 0$, the set of ε-critical values of q on B, $\Delta(q, \varepsilon)$, can be covered by $C(q)$ intervals of length ε, i.e. $M(\varepsilon, \Delta(q, \varepsilon)) \leqslant C(q)$, with $C(q)$ depending only on q and not on ε.* (For example, by Theorem 3.2 above, polynomials q of degree d, depending on n variables, satisfy this property with $C(q) \sim d^n$.)

Definition 8.1. For any f—a C^1-function on B, the Q-complexity $\sigma_Q(f, \varepsilon)$ is defined as

$$\sigma_Q(f, \varepsilon) = \inf_{q \in Q, |f - q|_{C^1} \leqslant \varepsilon} 3C(q).$$

Theorem 8.2. *For any* $\varepsilon > 0$,

$$M(\varepsilon, \Delta(f)) \leqslant \sigma_Q(f, \varepsilon).$$

The proof is identical to the proof of Theorem 3.6.

The main difficulty in the application of this result to specific functions on infinite-dimensional spaces consists of a choice of the approximating class Q. Below we give some examples where Q can be chosen in a natural way.

As the first choice we shall take Q consisting of polynomials, depending on a finite number of variables, or more accurately, on a finite number of linear functionals on V. We have to show that condition $(*)$ above is satisfied.

Proposition 8.3. *Let V be a Banach space and let $\ell_1, \dots, \ell_n$ be linear functions on V. Let $p(x_1, \dots, x_n)$ be a polynomial of degree d. Then for a function $\tilde{p} : B \to \mathbb{R}$, $\tilde{p}(v) = p(\ell_1(v), \dots, \ell_n(v))$, $C(\tilde{p}) \leqslant Kn \cdot (2d)^n$.*

Here $B \subseteq V$ denotes the unit ball in V, and $C(\tilde{p})$ is the minimal number of ε-intervals, covering the set of ε-critical values of $\tilde{p}$ on B. For $V = \ell^2$ and $\ell_1, \dots, \ell_n$ orthonormal, the factor n in the above expression can be omitted.

Proof. Consider the mapping $L : V \to \mathbb{R}^n$, $L(v) = (\ell_1(v), \dots, \ell_n(v))$. It is well known that there is a Euclidean structure on $\mathbb{R}^n$, such that $B^n_{1/\sqrt{n}} \subseteq L(B) \subseteq B^n_{\sqrt{n}}$, where B^n_r is the ball of radius r in this structure. Therefore, if $v \in B$ is an ε-critical point of $\tilde{p} = p \circ L$, then $L(v) \in \mathbb{R}^n$ is a $\sqrt{n}\varepsilon$-critical point of p, and this point belongs to the ball $B^n_{\sqrt{n}}$ (in the new Euclidean structure). Hence $\Delta(\tilde{p}, \varepsilon) \subseteq \Delta(p, \sqrt{n}\varepsilon)$, with p restricted to the ball $B^n_{\sqrt{n}}$, and the required result follows by Theorem 3.1. For $V = \ell^2$ and $\ell_1, \dots, \ell_n$ orthonormal, the coefficient $\sqrt{n}$ in the above inequality can be omitted. $\square$

Remark. The statement of Proposition 8.3 does not assume any restrictions on the norm of the linear functions $\ell_1, \dots, \ell_n$, as well as on the norm of the polynomial p. To provide this invariance, the proof has to appeal to the result on the approximation of any convex set by the Euclidean ball, as well as to the fact that no restrictions on the polynomial are assumed in the "polynomial Sard theorem" 3.2. For the purposes of the applications below it would be enough to consider only uniformly bounded linear functions on V.

Example 1 (Functions on ℓ^2). Consider the functions $f : \ell^2 \to \mathbb{R}$ of the form

$$f(x) = \sum_{i=1}^{\infty} \alpha_i p_i(x_1, \dots, x_i), \tag{8.2}$$

where $p_i(x_1, \dots, x_i)$ is a polynomial of degree d_i, such that $\max |p_i| \leqslant 1$ on the unit ball in $\mathbb{R}^i$, $i = 1, 2, \dots$

By the Markov inequality (see [Achi]), $\max ||dp_i|| \leqslant d_i^2$ on the unit ball in $\mathbb{R}^i$. Hence if α_i satisfy $\sum_{i=1}^{\infty} \alpha_i d_i^2 < \infty$, f is continuously differentiable on the unit ball $B \subseteq \ell^2$.

For f as above we define $\eta(f, \varepsilon)$ as follows: $\eta(f, \varepsilon) = [2 \max(d_1, \dots, d_{N(\varepsilon)})]^{N(\varepsilon)}$, where $N(\varepsilon)$ is the smallest natural N, for which the sum $\sum_{i=N+1}^{\infty} \alpha_i d_i^2 \leqslant \frac{1}{2}\varepsilon$.

Theorem 8.4. *For f defined by* (8.2), $\sigma_Q(f, \varepsilon) \leqslant \eta(f, \varepsilon)$. *In particular, for any* $\varepsilon > 0$, $M(\varepsilon, \Delta(f)) \leqslant \eta(f, \varepsilon)$.

Proof. We have for $N = N(\varepsilon)$, $|f - P_N|_{C^1} \leqslant \varepsilon$, where $P_N = \sum_{i=1}^{N} \alpha_i p_i$. But P_N is a polynomial of the variables $x_1, \dots, x_N$ in ℓ^2 and of the degree $\max(d_1, \dots, d_N)$. Hence, by Proposition 8.2, $C(P_N) \leqslant \eta(f, \varepsilon)$, and the result follows by Theorem 8.1. □

In particular, if all the degrees d_i are equal to d, and α_i form the geometric sequence $\alpha_i = \left(\frac{1}{q}\right)^i = \alpha^i$, $\alpha < 1$, we have the following: $\sum_{i=N+1}^{\infty} \alpha^i d^2 = d^2 \alpha^{N+1}/(1 - \alpha)$. For this to be smaller than or equal to ε we get $N = [\log_\alpha((1 - \alpha)\varepsilon d^2)]$, and $\eta(f, \varepsilon) \leqslant (2d)^{\log_\alpha((1-\alpha)\varepsilon d^2)} = K(\alpha, d)\left(\frac{1}{\varepsilon}\right)^{\log_q(2d)}$. Thus we have the following corollary:

Corollary 8.5. *For* $f = \sum_{i=1}^{\infty} \left(\frac{1}{q}\right)^i p_i(x_1, \dots, x_i)$ *with* $\deg p_i = d$, $|p_i| \leqslant 1$ *on the unit ball, and* $q > 1$, *we have* $\sigma_Q(f, \varepsilon) \leqslant K(q, d)\left(\frac{1}{\varepsilon}\right)^{\log_q(2d)}$. *In particular, for* $q > 2d$, f *satisfies the Sard theorem* (*i.e.*, $m(\Delta(f)) = 0$).

Returning to the Kupka example above, we see that it occurs exactly on the boundary: by Corollary 8.5, for any $q > 6$, functions $f = \sum_{i=1}^{\infty} \frac{1}{q^i} p_i(x, \dots, x_i)$ with $\deg p_i = 3$, satisfy the Sard theorem. (In the specific example above it is enough to take $q > 2$.) By a formal analogy we can say that the complexity of the function $f : \ell^2 \to \mathbb{R}$, $f = \Sigma \left(\frac{1}{q}\right)^i p_i(x_1, \dots, x_i)$, $\deg p_i = d$, is the same as the complexity of C^k-functions $g : B^n \to \mathbb{R}$, if $\frac{n}{k-1} = \log_q(2d) = \beta$. In particular, the sequences of the form $1, 1/2^s, 1/3^s, \dots, 1/k^s, \dots$, may appear among the critical values of both f and g only if $1/(s-1) \leqslant \beta$.

Pursuing this formal analogy a little bit further, we find that the functions $f : \ell^2 \to \mathbb{R}$, $f(x) = \sum_{i=1}^{\infty} \alpha_i p_i(x_1, \dots, x_i)$, have the complexity of the usual analytic functions, if the coefficients α_i tend to zero very fast: $\alpha_i \sim \left(\frac{1}{q}\right)^{b^i}$, where $q > 1$, $b > 1$. Indeed, proceeding exactly as above, we find, that $|f - P_N|_{C^1} \leqslant \varepsilon$, if $\left(\frac{1}{q}\right)^{b^N} \leqslant \varepsilon$, or for $N \geqslant \log\left(\frac{\ln(1/\varepsilon)}{\ln q}\right)$, and for $\deg p_i \equiv d$, we get the following:

Corollary 8.6. *For* $f=\sum_{i=1}^{\infty}\alpha_i p_i(x_1,\dots,x_i)$, *with* $|p_i|\leqslant 1$ *on the unit ball*, $\deg p_i=d$ *and* $\alpha_i=\left(\frac{1}{q}\right)^{b^i}$, $q>1$, $b>1$, *we have* $\sigma_Q(f,\varepsilon)\leqslant c(b,q,d)\left(\frac{\ln(1/\varepsilon)}{\ln q}\right)^{\log_b(2d)}$.

Comparing this expression with the remark after Theorem 2.3 above, we can say formally, that for such f the number $\log_b(2d)$ plays the role of the dimension, while the number $\ln q$ plays the role of the size h of a complex neighborhood, to which the analytic function extends holomorphically. In both cases of Corollaries 8.5 and 8.6 the examples show that these complexity estimates are essentially sharp.

Example 2 (Functions on $C^k[0,1]$). Our second example is based on the same approximating class Q of polynomials in finite number of linear functionals. As the underlying Banach spaces we take the spaces $C^k[0,1]$ of k times continuously differentiable functions on $[0,1]$ with the standard C^k-norm. In fact, we consider the same function ϕ (given by expression (8.3) below) on all the chain of the spaces $C^0\subset C^1\subset\cdots\subset C^k\subset\cdots$, and study the dependence of the complexity of ϕ on k.

Fix some $q>1$ and consider the sequence of points $x_i=(1/q)^i\in[0,1]$. Let $\varphi_i(x)$ be polynomials of one variable of degree d, such that $|\varphi_i(x)|\leqslant 1$ for $|x|\leqslant 1$. Now let B be the unit ball in $C^0[0,1]$, consisting of all the continuous functions u on $[0,1]$, with $\max_{[0,1]}|u|\leqslant 1$. Define $\phi: B\to\mathbb{R}$ by

$$\phi(u)=\sum_{i=1}^{\infty}\alpha_i\varphi_i(u(x_i)),\quad u\in C^0[0,1]. \tag{8.3}$$

For the derivative of ϕ in the direction of a function $v\in C^0[0,1]$, we have

$$\frac{d}{dv}\phi(u)=\frac{d}{dt}\phi(u+tv)|_{t=0}=\sum_{i=1}^{\infty}\alpha_i\varphi_i'(u(x_i))v(x_i). \tag{8.4}$$

In particular, we see as above that if $\sum_{i=1}^{\infty}|\alpha_i|d_i^2<\infty$, then ϕ is (Frechet) continuously differentiable.

The above formula shows that u is a critical point of ϕ (i.e. $\frac{d}{dv}\phi(u)=0$ for any v) if and only if $\alpha_i\varphi_i'(u(x_i))=0$ for any i, or, equivalently, $u(x_i)$ is a critical point of φ_i for each i with $\alpha_i\neq 0$.

We can easily produce examples of functions of the form (8.3) violating Sard theorem on the spaces $C^k[0,1]$. Let φ be the same polynomial of degree 3 as in Kupka's example: $\varphi(0)=\varphi'(0)=0$, $\varphi(1)=1$ and $\varphi'(1)=0$. Consider a special case Ψ of the function ϕ, defined by

$$\Psi_r(u)=\sum_{i=1}^{\infty}\left(\frac{1}{2}\right)^i\varphi(r^i u(x_i)), \tag{8.5}$$

where $r>1$ is fixed.

Critical points of Ψ_r on $C^0[0,1]$ are exactly those continuous functions u, for which $u(x_i)=a_i(1/r^i)$, $a_i=0,1$. Since $r>1$, such continuous u exist for any choice of

a_i. The value $\Psi_r(u)$ for any such u is equal to $\sum_{i=1}^{\infty}(a_i/2^i)$, and hence these critical values cover all the interval $[0, 1]$.

Now one can easily see that for our choice of $x_i = (1/q)^i$, for any choice of $a_i = 0, 1$, the functions u, satisfying $u(x_i) = a_i(1/r^i)$ can be chosen to be C^k on $[0, 1]$, if $\log r / \log q \geqslant k$. (Indeed, $u(x_i) = a_i(1/r^i)$ is of order x_i^μ, $\mu = \frac{\log r}{\log q}$.) Thus the functions Ψ_r given by (8.5), with growing r, violate the Sard theorem on subspaces $C^k[0, 1]$ of $C^0[0, 1]$, for $r > q^k$.

Let us now estimate the Q-complexity of a (more general) function ϕ, defined by (8.3), on various $C^k[0, 1]$. Let $\varepsilon > 0$ be given. To ε-approximate ϕ by polynomials of a finite number of variables, we proceed as follows: subdivide $[0, 1]$ into two parts: $[0, \delta]$ and $[\delta, 1]$, $1 > \delta > 0$. The approximating polynomial P^δ is completely defined by the parameter δ as follows:

We have $x_i = (1/q)^i \in [0, \delta]$ for $i \geqslant [\log_q(1/\delta)] + 1 = N(\delta)$. Define $P_1^\delta(u)$ by $P_1^\delta(u) = \sum_{i=1}^{N(\delta)-1} \alpha_i \varphi_i(u(x_i))$. Let $\tilde{u}$ be the Taylor polynomial of $u \in C^k[0, 1]$, of degree $k-1$ at the point $0 \in [0, 1]$. We define $P_2^\delta(u)$ as $P_2^\delta(u) = \sum_{i=N(\delta)}^{\infty} \alpha_i \varphi_i(\tilde{u}(x_i))$. Finally, we put $P^\delta(u) = P_1^\delta(u) + P_2^\delta(u)$. We notice that $P_2^\delta(u)$ is in fact a polynomial of degree d in k linear functionals on $C^k([0, 1])$: $u(0), u'(0), \ldots, u^{(k-1)}(0)$. Thus P^δ is a polynomial of degree d in $N(d) + k - 1$ linear functionals $u(x_1), u(x_2), \ldots, u(x_{N(d)-1}), u(0), u'(0), \ldots, u^{(k-1)}(0)$ on $C^k[0, 1]$.

We have $|\phi - P^\delta|_{C^1} \leqslant \sup_{||u|| \leqslant 1} (|\phi(u) - P^\delta(u)| + ||d\phi(u) - dP^\delta(u)||)$.

Since $\phi(u) - P^\delta(u) = \sum_{i=N(\delta)}^{\infty} \alpha_i(\varphi_i(u(x_i)) - \varphi_i(\tilde{u}(x_i)))$ and since, by the Taylor formula,

$|u(x) - \tilde{u}(x)| \leqslant \frac{1}{k!}\delta^k$ for $x \in [0, \delta]$, we get $|\phi(u) - P^\delta(u)| \leqslant \frac{d^2}{k!}\delta^k \sum_{i=N(\delta)}^{\infty} \alpha_i$. (By the Markov inequality, $\varphi_i' \leqslant d^2$ on $[0, 1]$. Remind also that u belongs to the unit ball in $C^k[0, 1]$.)

For the derivative at u in the direction of the function $v \in C^k[0, 1]$, $||v|| \leqslant 1$, we get

$$|d\phi(u)(v) - dP^\delta(u)(v)| \leqslant \sum_{i=N(\delta)}^{\infty} \alpha_i |\varphi_i'(u(x_i)) \cdot v(x_i) - \varphi_i'(\tilde{u}(x_i))\tilde{v}(x_i)|,$$

where as above $\tilde{u}$ and $\tilde{v}$ are the Taylor polynomials of u and v, respectively.

Therefore,

$$\begin{aligned} &|\varphi_i'(y(x_i))v(x_i) - \varphi_i'(\tilde{u}(x_i))\tilde{v}(x_i)| \\ &\quad \leqslant |\varphi_i'(u(x_i)) - \varphi_i'(\tilde{u}(x_i))| \cdot |v(x_i)| + |\varphi_i'(\tilde{u}(x_i))| \cdot |\tilde{v}(x_i) - v(x_i)| \\ &\quad \leqslant d^2(d-1)^2 \frac{1}{k!}\delta^k + d^2 \frac{1}{k!}\delta^k \leqslant \frac{d^4}{k!}\delta^k, \end{aligned}$$

and $||d\phi - dP^\delta|| \leqslant \frac{d^4}{k!} \sum_{i=N(\delta)}^{\infty} \alpha_i$.

Finally, we get $|\phi - P^\delta|_{C^1} \leqslant \frac{2d^4}{k!}\delta^k \sum_{i=N(\delta)}^{\infty} \alpha_i \stackrel{\text{def}}{=} \eta(\delta)$.

Clearly, the accuracy $\eta(\delta)$ of the approximation of ϕ by P^{δ} increases as δ decreases to zero.

Now, let $\delta(\varepsilon)$ be the maximal δ such that $\eta(\delta) \leqslant \varepsilon$ and let $N(\varepsilon) = N(\delta(\varepsilon)) + k - 1$. Exactly as above, we get $\sigma_Q(\phi, \varepsilon) < 4(2d)^{N(\varepsilon)} \cdot N(\varepsilon) \stackrel{\text{def}}{=} \eta(\phi, \varepsilon)$. Therefore we have proved the following:

Theorem 8.7. *For ϕ as above, the Q-complexity of ϕ satisfies $\sigma_Q(\phi, \varepsilon) \leqslant \eta(\phi, \varepsilon)$.*

To give explicit bound, assume, for example, that $\alpha_i = (1/s)^i$. Then

$$\sum_{i=N(\delta)}^{\infty} \alpha_i \sim (1/s)^{N(\delta)} \sim (1/s)^{\log_q(1/\delta)} = \delta^{\log_q s}.$$

Hence $\eta(\delta) \sim \delta^{k+\log_q s}$, $\delta(\varepsilon) \sim \varepsilon^{1/(k+\log_q s)}$, $N(\varepsilon) \sim \frac{\log_q(1/\varepsilon)}{k+\log_q s}$, and finally $\eta(\Phi, \varepsilon) = 4(2d)^{N(\varepsilon)} \cdot N(\varepsilon) \sim (1/\varepsilon)^{(\log_q(2d)/k+\log_q s)} \cdot N(\varepsilon)$. We have therefore:

Corollary 8.8. *For $\phi : C^k[0,1] \to \mathrm{R}$, defined by $\phi(u) = \sum_{i=1}^{\infty} (1/s)^i \varphi_i(u(x_i))$, where $x_i = (1/q)^i$, $s, q > 1$, and where φ_i are polynomials of one variable of degree d with $|\varphi_i(x)| \leqslant 1$, $x \in [0,1]$, we have*

$$\sigma_Q(\phi, \varepsilon) \leqslant K(s, q, k, d)(1/\varepsilon)^{(\log_q(2d)/k+\log_q s)} \cdot \log_q(1/\varepsilon).$$

Thus the same function ϕ considered on the spaces $C^k[0,1]$ for various k, has various complexities: the higher is k, the simpler is ϕ.

Corollary 8.9. *The measure of the set of critical values $\Delta(\phi)$ of the function ϕ, restricted to $C^k[0,1]$, is zero, provided that $k > \log_q(2d/s)$.*

Proof. We just solve $\frac{\log_q(2d)}{k+\log_q s} < 1$.

Let us now return to the functions Ψ_r, violating the Sard theorem on $C^k[0,1]$ for $r > q^k$. To apply Corollary 8.9 we have first to renormalize polynomials $\varphi(r^i u)$ in order to have them bounded by 1 for $|u| \leqslant 1$. We can rewrite (8.5) in the following form:

$$\Psi_r(u) = \sum_{i=1}^{\infty} \left(\frac{r^3}{2}\right)^i \frac{1}{r^{3i}} \varphi(r^i u(x_i)) = \sum_{i=1}^{\infty} \left(\frac{r^3}{2}\right)^i \varphi_i(u(x_i)). \tag{8.6}$$

(In particular, for (8.6) to converge on B, we shall assume $1 < r \leqslant \sqrt[3]{2}$.) Now, in (8.6), $s = 2/r^3$, and Corollary 8.9 gives

$$m(\Delta(\Psi_r)) = 0 \quad \text{if } k > \log_q(3r^3),$$

or $r \leqslant (1/3)q^{(k/3)}$. As it was shown above, $\Delta(\Psi_r)$ is $[0,1]$ for $r > q^k$. More accurate computations (and considering more general sequences of points x_i in the construction of the functions ϕ and Ψ_r), allow one to show that the bounds above are rather accurate. $\square$

Example 3 (Control problems). In this section we give another example of a function on an infinite-dimensional space: the so called control-to-state mapping in a control problem $\dot{x} = f(x, u)$. The above approach works for this function, with the approximating class Q, given by polynomial control problems $\dot{x} = p(x, u)$, p—a polynomial. As usual, we discuss for this important class of nonlinear mappings the semialgebraic complexity (and, in particular, the problem of the validity of Sard's theorem and of its "quantitative" generalizations).

We consider a nonlinear finite-dimensional control problems of the type

$$\dot{x}(t) = f(x(t), u(t)), x(0) = x_0, \quad t \in [0, T], \tag{8.7}$$

where $x(t)$ is the state, and $u(t)$ is the control, at time t.

Input-to-state mapping J_f of (8.7) associates to each control $\tilde{u}$ the state $J_f(\tilde{u})$, to which $\tilde{u}$ steers the system from the initial state x_0 in time T. Thus J_f is a mapping of an infinite-dimensional space of the allowed controls $u(t)$ into the finite-dimensional state space of the system.

Mappings J_f for nonlinear f are known to be complicated. However, the question of validity of Sard's theorem for these mappings is important from both the theoretical and computational points of view (see [Bri-Yom1,Bri-Yom2,Bri-Yom3,Bri-Yom4,Sus]).

In contrast with Examples 1 and 2, in our third example it is not immediately clear how to approximate J_f by "polynomials in a finite number of variables". However the results of [Bri-Yom1,Bri-Yom2,Bri-Yom3,Bri-Yom4] suggest another natural class Q of "simple" approximants: the input-to-state mappings J_f of the control problems (8.7) with the right-hand side f a polynomial.

We remind shortly some of the results of [Bri-Yom2]. Let us assume x and u to be one-dimensional. (Our methods work also in a multi-dimensional situation.)

Differential $DJ_f(u)(v)$ is given by the solution $z(T)$ of the linearized equation (8.8) along the trajectory $(x(t), u(t))$:

$$\dot{z}(t) = f_z(x(t), u(t))z(t) + f_u(x(t), u(t))v(t), z(0) = 0. \tag{8.8}$$

In particular, a control $u(t)$ is critical for J_f if and only if $f_u(x(t), u(t)) \equiv 0$.

As we assume f to be a polynomial, this last equation $f_u(x, u) = 0$ defines an algebraic curve Y in the plane (x, u). It allows one to express u as a (generally multivalued) function $u(x)$ of x.

Choosing a certain univalued branch $u(x)$ of this multivalued function and substituting this expression into the original equation (8.7) we get an ordinary differential equation

$$\dot{x} = f(x, u(x)), x(0) = x_0, \quad t \in [0, T], \tag{8.9}$$

whose solution is uniquely defined on a certain subinterval of the interval of regularity of the chosen branch.

Hence assuming that the control $u(t)$ is critical (and that it is continuous, i.e. it does not jump from one branch of the algebraic curve Y to another at the points where this jump is nonzero), we get only a finite number of possibilities for the control u and for the solution x: at any double (multiple) point of the algebraic curve Y the control can switch from one its branch to another. Clearly, the total number of such choices for u is bounded through the degree of the polynomial f.

This simple consideration shows that for the polynomial control problem (8.7) the number of critical values of the input-to-state mapping J_f (on the space of continuous controls) is finite, and bounded through the degree of the polynomial f.

However, in order to use the input-to-state mapping J_f (for f polynomials) as "simple" approximant, we have, according to the condition $(*)$ on Q stated in the beginning of this section, to get bounds not only on the critical but also on the near-critical values of J_f.

Let us consider near-critical controls in (8.7). If the differential of the input-to-state mapping J_f is small, the differential equation (8.8) becomes a differential inequality, which leads to the requirement that the absolute value of $f_u(x,u)$ be small. This condition defines a semialgebraic set S in the plane. (All these objects of course depend on the parameter, measuring the size of the differential of J_f.)

Therefore, near-critical trajectories $(x(t),u(t))$ lie in S. The complement to S consists of a finite (and bounded through the degree of f) number of "islands" O_i. Let us assume that $x(t)$ is monotone in t on $[0,T]$. (If a near critical trajectory $x(t)$ "turns back" at a certain moment t_0, one can show that it remains near the turning point $x(t_0)$for the rest of the time. See [Bri-Yom2].) Then for each island O_i in the plane (x,u) the trajectory $(x(t),u(t))$ can pass either above or below O_i. Now two trajectories, that pass on the same side of each of the islands O_i, are "visible" one from another. Using special metric properties of semialgebraic sets, proved in [Bri-Yom1,Bri-Yom2,Bri-Yom3,Bri-Yom4] (see also [Yom-Com]), one can join these trajectories inside S by paths of controllable lengths, and to estimate the difference of the derivatives of $x(t)$. As a result, we get a differential inequality, which implies that the endpoints of the two trajectories as above must be close to one another.

The following result is obtained in [Bri-Yom2] by a detailed analysis in the above lines:

Denote W_K the set of K-Lipschitzian controls u on [0,1] with $|u(t)| \leqslant 1$, and fix L_p-norm on the control space, $p \geqslant 1$.

Theorem 8.10. *Assume* $x_0 = 0$ *in* (8.7). *Let* $f(x,u)$ *be a polynomial of degree* d, *satisfying* $|f(x,u)| \leqslant 1$ *for* $|x| \leqslant 1$, $|u| \leqslant 1$. *Then for any* $1 \geqslant \gamma \geqslant 0$ *the set of* γ*-critical values of* J_f *on* W_K *can be covered by* $N(d) = 2^{4(d+1)^2}$ *intervals of length* $\delta = (qK)^{1/q}\gamma^{q/q+1}$. *Here* $1/p + 1/q = 1$.

In particular, for $p = 1$ and $q = \infty$ we get $\delta = \gamma$. Thus, the quantitative Sard theorem is valid for the polynomial control problems as above, and we get the required property $(*)$ of the approximating class Q, consisting of their input-to-state functions.

Now we apply the approach of Section 8 in order to extend the result to more complicated right-hand sides than polynomials. Notice, however, that since the growth of the estimate of Theorem 8.10 in d is very fast, a very high regularity of f will be necessary to guaranty the validity of the Sard theorem.

In what follows, we give only a sketch of the proof of the "semialgebraic complexity" result for J_f. We plan to present separately additional results in this direction, as well as the detailed proofs.

If a polynomial p ε-approximates f in an appropriate norm, then it is easy to see that J_p ε-approximates J_f in the L_1-norm on the control space.

According to Definition 8.1, the Q-complexity $\sigma_Q(J_f, \varepsilon)$ is defined as

$$\sigma_Q(J_f, \varepsilon) = \inf_{J_p \in Q, |J_f - J_q|_{C^1} \leqslant \varepsilon} 3C(J_p).$$

By Theorem 8.10, $C(J_p) \leqslant N(d)$, where $d = \deg p$, and taking into account the remark above, we get

Theorem 8.11. *For any* $\varepsilon > 0$,

$$\sigma_Q(J_f, \varepsilon) \sim \inf_{\deg p = d, |f-p|_{C^1} \leqslant \varepsilon} 3N(d).$$

Now for f, for example, in C^k, the degree d of the ε-approximating polynomial p is of order $\left(\frac{1}{\varepsilon}\right)^{\frac{1}{k}}$, and we get

$$\sigma_Q(J_f, \varepsilon) \sim 2^{4\left(\frac{1}{\varepsilon}\right)^{\frac{2}{k}}}.$$

Of course, the corresponding bound for the critical values inside the real line does not provide any nontrivial restriction. A simple computation shows that in order to get a complexity $\sigma_Q(J_f, \varepsilon)$ growing as $\left(\frac{1}{\varepsilon}\right)^{\alpha}$ with $\alpha < 1$, we must have the degree d of the ε-approximating polynomial p of f to be of the order of $\sqrt{\alpha \log\left(\frac{1}{\varepsilon}\right)}$ $((**))$. Equivalently, the best approximation of f by polynomials of degree d must be of the order of $\left(\frac{1}{2}\right)^{4d^2/\alpha}$. So only entire functions with very rapidly decreasing Taylor coefficients a_i satisfy this condition. In particular, this is true if $a_i \sim \left(\frac{1}{2}\right)^{4i^2/\alpha}$.

Thus we get

Theorem 8.12. *For any f satisfying* $(**)$ *and for any* $\varepsilon > 0$,

$$M(\varepsilon, \Delta(f)) \leqslant K\left(\frac{1}{\varepsilon}\right)^{\alpha}.$$

The same bound is true for an ε-neighborhood of $\Delta(f)$. In particular, no sufficiently long finite part of the sequence $1, \left(\frac{1}{2}\right)^{\beta}, \left(\frac{1}{3}\right)^{\beta}, \ldots$ can be in $\Delta(f)$, if $\beta < \frac{1}{\alpha} - 1$.

The description given above for critical controls of a polynomial one-dimensional control problem (8.7) remains valid also for f analytic. Still these controls go along the branches of an analytic curve $f_u(x,y) = 0$, switching from one branch to another only at double points of this curve. Hence the number of continuous critical controls of an analytic control problem (8.7) is finite, as well as the number of its critical values. However, the bounds on the geometry of the critical and near critical values, and in particular, on their ε-entropy, obtained by our method, remain nontrivial, since they provide strong restrictions on the number of points in $\Delta(f)$ and on their mutual position. We do not formulate here the results on the near-critical values of (8.7), restricting ourselves in Theorem 8.12 to the critical values only.

There is another direction, providing nontrivial examples of a validity of the Sard theorem for the control problem (8.7). Indeed, to conclude that the Lebesgue measure of $\Delta(f)$ is zero, it is enough to get the bound of Theorem 8.12 only for a certain sequence of the values of ε, i.e. only for a subsequence of the degrees of the approximating polynomials. Hence, the considerations of Section 6 work in our current context, and we get the Sard theorem for J_f with f certain quasianalytic functions in the sense of S. Bernstein.

Using the same considerations as above, but with an infinite number of the "islands" O_i, one can construct control problems of the above form with f infinitely smooth and with critical values of J_f covering the whole interval.

9. Some concluding remarks

9.1. *Average topological complexity of fibers*

Let $f: M^m \to N^n$ be a C^k mapping of compact manifolds. The following implication of the usual Sard theorem is by far the most frequently used in differential topology: for almost any y in N the fiber $f^{-1}(y)$ is a compact smooth submanifold of M.

A natural question is then what is a typical topological complexity of such a fiber (in particular, how many connected components may it have)? The usual Sard theorem gives no information of this type. On the other hand, the Quantitative Sard theorem (Theorem 3.1, stated above) allows us to give explicit upper bounds for an average of Betti numbers of the fibers.

Let for y in N, $B_i(y)$ denote the ith Betti number of the fiber $f^{-1}(y)$. We assume that the smoothness k of f is greater than $s = n - m + 1$. Then by the (usual) Sard theorem $B_i(y)$ is finite almost everywhere in N.

Theorem 9.1. *For any q between zero and $k - s/n$ the average of $(B_i(y))^q$ over N is finite.*

The proof of this theorem is given in [Yom3]. It is based on an estimate of the average distance from a point in N to the set of critical values of f. This estimate is provided by the Quantitative Sard theorem.

Many additional results of the same spirit are obtained in [Yom3]: existence of "simple" fibers in any subset of N of a given positive measure, average bounds on the volume of the fibers, etc. For M and N Euclidean balls, explicit bounds for all the above quantities are given in [Yom3] in terms of the bounds on the derivatives of f.

Semialgebraic complexity works perfectly well in this situation. The result of Theorem 9.1 remains valid for f only twice differentiable, but with a complexity of a C^k-function. With some more effort one can get a version of Theorem 9.1 for C^1 and even for Lipschitzian functions f of a low semialgebraic complexity. We plan to present these results separately.

Also here, much more information can be presumably extracted by similar methods: estimates for "bounded triangulations" of the fibers, more delicate estimates of the geometry, including upper bounds for curvatures, estimates for the spectrum of certain differential operators on these fibers.

9.2. *Other examples of simple functions*

We did not discuss here many additional natural examples of functions with prescribed semialgebraic complexity. Some of them, such as infinite sums of semialgebraic functions, functions with different smoothness in different groups of variables, functions allowing for a "separation of variables" ($F(x,y) = f(x) + g(y), F(x,y) = f(x) \cdot g(y)$), etc.; such functions can be treated also as special cases of compositions) are shortly discussed in [Yom18].

9.3. *Simple functions arising in dynamics*

In dynamics starting with smooth or analytic data, we often come naturally to highly non-regular objects. For example, such are in most cases homeomorphisms, conjugating two regular systems. It would be very interesting to investigate complexity of such objects. In some cases (as, for example, for a conjugating homeomorphism of the circle, which linearizes an analytic automorphism with a "bad" rotation number) a lacunary sequence of very fast analytic approximations exists. However, we expect that significant modifications of the notion of a semialgebraic complexity, as developed above, will be necessary.

9.4. *Complexity of iterations*

As it was mentioned above, the bounds on the complexity growth in iterations of smooth or analytic mappings, obtained in [Yom13,Yom14,Yom15,Yom16], are not valid for general nonsmooth functions with low semialgebraic complexity. This is because if we allow "semialgebraic partitions", their local "combinatorial" complexity can blow up exponentially in iterations. This problem disappears, however, for "lacunary series" considered in Section 6 above. Moreover, the behavior of "lacunary series" under compositions is quite controllable (if f is well approximated by polynomials p_j of degrees d_j, $f \circ f$ is well approximated by polynomials $p_j \circ p_j$ of degrees $d_j^2 \ldots$) and resembles the behavior of smooth functions. Consequently, the following question seems to be very important for a better understanding of the nature of complexity of functions:

Can the bounds on the local complexity growth in iterations, obtained in [Yom13,Yom14,Yom15,Yom16], *be extended to the "lacunary series"*?

9.5. Relation to the fixed point theorem

There is a conjectured relation between the validity of the Sard theorem for a certain operator on a Banach space, and the validity of the fixed point theorem for this operator. Indeed, it is well known that one can prove fixed point theorem using smooth approximations and the Sard theorem (see a beautiful presentation of this proof in [Mil]). It would be very interesting to investigate this possible relation in the situations, where neither topological fixed point methods nor analytic methods (like KAM) work. Many such situations arise in nonlinear dynamics and differential equations, and the complexity of the operators involved in some cases presumably can be estimated.

Acknowledgments

The author thanks the referee for a number of important improvements and corrections.

References

[Achi] N.I. Achieser, Theory of Approximation, Dover Publications, New York, 1992.

[Al] A.D. Alexandrov, Almost everywhere existence of the second differential of a convex function and some properties of convex functions connected with it, Uchenye Zap. Math. Ser. Leningr. Univ. 6 (1939) 3–35.

[AZ] A.D. Alexandrov, V.A. Zalgaller, Intrinsic Geometry of Surfaces, AMS, Providence, RI, 1967.

[Ber1] S. Bernstein, Sur une propriété des polynomes, Proc. Kharkov Math. Soc. Ser. 2 (14) (1913) 1–6.

[Ber2] S. Bernstein, Sur la définition et les propriétés des fonctions analytiques d'une variable réelle, Math. Annalen 75 (1914) 449–468.

[Ber3] S. Bernstein, Sur les fonctions quasianalytiques, C.R. Paris 177 (1923) 937–939.

[Bic-Yom] E. Bichuch, Y. Yomdin, High order Taylor discretization for numerical solution of parabolic PDE's, preprint, 1995.

[Boc-Cos-Roy] J. Bochnak, M. Coste, M.F. Roy, Real algebraic geometry, Ergeb. Math. Grenzgeb. 36 (3) (1998) (translated from the 1987 French original. Revised by the authors).

[BF] T. Bonnesen, W. Fenchel, Teorie der konvexen Körper, Springer, Berlin, 1934.

[Bri-Yom1] M. Briskin, Y. Yomdin, Vertices of reachable sets in nonlinear control problems, Proceedings of the 29th CDC, Honolulu, Hawaii, December, 1990, pp. 2815–2816.

[Bri-Yom2] M. Briskin, Y. Yomdin, Critical and near-critical values in polynomial control problems, I: one-dimensional case, Israel J. Math. 78 (1992) 257–280.

[Bri-Yom3] M. Briskin, Y. Yomdin, Extremal and near-extremal trajectories of rank zero in nonlinear control problems, in: A. Ioffe, M. Marcus, S. Reich (Eds.), Optimization and Nonlinear Analysis, Pitman Research Notes in Mathematics, Series, Vol. 244, Pitman, London, 1992, pp. 53–63.

[Bri-Yom4] M. Briskin, Y. Yomdin, Semialgebraic geometry of polynomial control problems, in: F. Eyssette, A. Galligo (Eds.), Computational Algebraic Geometry, Birkhaüser, Basel, 1993, pp. 21–28.

[Bri-Eli-Yom] M. Briskin, Y. Elichai, Y. Yomdin, How can Singularity theory help in Image processing, in: M. Gromov, A. Carbone (Eds.), Pattern Formation in Biology, Vision and Dynamics, World Scientific, Singapore, 2000, pp. 392–423.

[Bry] L.N. Bryzgalova, Singularities of a maximum of a function depending on parameters, Funct. Anal. Appl. 11 (1977) 49–50.

[Cla1] F.H. Clarke, Generalized gradients and applications, Trans. Amer. Math. Soc. 205 (1975) 247–262.

[Cla2] F.H. Clarke, On the inverse function theorem, Pacific J. Math. 64 (1976) 97–102.

[Dol1] E.P. Dolgenko, The rate of a rational approximation and properties of functions, Mat. Sb. 56 (1962) 403–432.

[Dol2] E.P. Dolgenko, On the properties of functions of several variables, well approximable by rational functions, Izv. Acad. Nauk SSSR 26 (1962) 641–652.

[Eli-Yom1] Y. Elihai, Y. Yomdin, Flexible high-order discretization, preprint, 1989.

[Eli-Yom2] Y. Elihai, Y. Yomdin, Global motion planning algorithm, based on a high order discretization and on hierarchies of singularities, Proceedings of the 28th CDC, Tampa, Florida, 1989, pp. 1173–1174.

[Eli-Yom3] Y. Elihai, Y. Yomdin, Flexible high order discretization of geometric data for global motion planning, Theoretical Computer Science A, Vol. 157, Elsevier Science B.V., Amsterdam, 1996, pp. 53–77.

[Eli-Yom4] Y. Elihai, Y. Yomdin, Numerical inversion of plane to plane mappings, based on a high order discretization and on hierarchies of singularities, in preparation.

[Eli-Yom5] Y. Elihai, Y. Yomdin, Normal forms representation: a technology for image compression, Image and Video Processing, Vol. 1903, SPIE, 1993, pp. 204–214.

[Gol-Gui] V. Golubitski, V. Guillemin, Stable Mappings and their Singularities, Graduate Texts in Mathematics, Vol. 14, Springer-Verlag, Berlin, 1973.

[Iva1] L.D. Ivanov, Variazii množhestv i funktsii (Russian), [Variations of sets and functions], in: A.G. Vituškin (Ed.), Izdat, Nauka, Moscow, 1975, p. 352pp.

[Iva2] L.D. Ivanov, The approximation of l-smooth functions by rational ones in the integral metric (Russian), Izv. Akad. Nauk SSSR Ser. Mat. 36 (1972) 234–239.

[Kho] A. Khovanskii, Fewnomials, AMS, Providence, RI, 1991.

[Ko-Se-Yom1] E. Kochavi, R. Segev, Y. Yomdin, Numerical solution of field problems by nonconforming Taylor discretization, Appl. Math. Modeling 15 (1991) 152–157.

[Ko-Se-Yom2] E. Kochavi, R. Segev, Y. Yomdin, Modified algorithms for nonconforming Taylor discretization, Comput. Struct. 49 (6) (1993) 969–979.

[Kol] A.N. Kolmogorov, On the representation of continuous functions of several variables by superpositions of continuous functions of fewer variables, DAN SSSR 108 (1956) 179–182 (AMS Transl. (2) 17 (1961), 369–373).

[Kol-Tih] A.N. Kolmogorov, V.M. Tihomirov, ε-entropy and ε-capacity of sets in functional space, Amer. Math. Soc. Transl. 17 (1961) 277–364.

[Kup] I. Kupka, Counterexample to the Morse–Sard theorem in the class of infinite-dimensional manifolds, Proc. AMS 16 (1965) 954–957.

[Lor1] G.G. Lorentz, Approximation of Functions, 2nd Edition, Chelsea Publishing Co., New York, 1986, 188pp.

[Lor2] G.G. Lorentz, Metric entropy and approximation, Bull. Amer. Math. Soc. 72 (1966) 903–937.

[Mat] V.I. Matov, Topological classification of germs of maximum and minimax functions generic families, Uspekhi Mat. Nauk 37 (4) (1982) 167–168.

[Mil] J. Milnor, Topology from the differentiable viewpoint, The University Press of Virginia, Charlottesville, 1969.

[Mor] A. Morse, The behavior of a function on its critical set, Ann. Math. 40 (1939) 62–70.

[Roc] R.T. Rockafeller, Favorable classes of Lipschitz continuous functions in subgradient optimization, in: E. Nurminski (Ed.), Progress in Nondifferentiable Optimization, Pergamon Press, New York, 1981.

[Roh1] A. Rohde, On Sard theorem for nonsmooth functions, Numer. Funct. Anal. Optim. 18 (9–10) (1997) 1023–1039.

[Roh2] A. Rohde, On the density of near-critical values of continuous selections and maximum functions. Approximation and Optimization in the Caribbean, II (Havana, 1993), Approx. Optim., Vol. 9, Lang, Frankfurt am Main, 1997, pp. 569–581.

[Roh3] A. Rohde, On the ε-entropy of nearly critical values, J. Approx. Theory 76 (2) (1994) 166–176.

[Sar1] A. Sard, The measure of the critical values of differentiable maps, Bull. Amer. Math. Soc. 48 (1942) 883–890.

[Sar2] A. Sard, Images of critical sets, Ann. Math. 68 (1958) 247–259.

[Sha-Yom] A. Shapiro, Y. Yomdin, On functions, representable as a difference of two convex functions, and necessary conditions in a constrained optimization, preprint, Beer-Sheva, 1981.

[Shu-Sma1] M. Shub, S. Smale, Computational complexity, On the geometry of polynomials and a theory of cost. I, Ann. Sci. éc. Norm. Sup. 4 Sér. t. 18 (1) (1985) 107–142.

[Shu-Sma2] M. Shub, S. Smale, Computational complexity: On the geometry of polynomials and a theory of cost. II, SIAM J. Comput. 15 (1) (1986) 145–161.

[Sma] S. Smale, An infinite dimensional version of Sard's theorem, Amer. J. Math. 87 (1965) 861–866.

[Sus] H.J. Sussman, Lie brackets, real analyticity and geometric control, Differential Geomet. Control Theory, Progr. Math., 27, Birkhauser Boston, Boston, MA, 1983, 1–116.

[Tan-Yom] A. Tannenbaum, Y. Yomdin, Robotic manipulators and the geometry of real semialgebraic sets, IEEE Trans. Robot. Automat. 3 (4) (1987) 301–307.

[Tih] V.M. Tihomirov, Kolmogorov's work on the ε-entropy of functional classes and superposition of functions, Russian Math. Surveys 18 (5) (1963) 51–87.

[Tri] C. Tricot, Douze définitions de la densité logarithmique (French. English summary) [Twelve definitions of logarithmic density], C. R. Acad. Sci. Paris Sér. I Math. 293 (11) (1981) 549–552.

[Vit1] A.G. Vitushkin, O mnogomernyh Variaziyah, Gostehisdat, Moscow, 1955.

[Vit2] A.G. Vitushkin, Ozenka sloznosti zadachi tabulirovaniya, Fizmatgiz, Moscow, 1959 (translation: Theory of the transmission and processing of information. Pergamon Press, 1961).

[Vit3] A.G. Vitushkin, On representation of functions by means of superposition and related topics, Enseignement Math. 23 (3–4) (1977) 255–320.

[War] H.E. Warren, Lower bounds for approximation by nonlinear manifolds, Trans. Amer. Math. Soc. 133 (1968) 167–178.

[Whi1] H. Whitney, A function not constant on a connected set of critical points, Duke Math. J. 1 (1935) 514–517.

[Whi2] H. Whitney, On singularities of mappings of Euclidean spaces I, Mappings of the plane into the plane, Ann. Math. 62 (1955) 374–410.

[Wie-Yom] Z. Wiener, Y. Yomdin, From formal numerical solutions of elliptic PDE's to the true ones, Math. Comput. 69 (229) (2000) 197–235.

[Yom1] Y. Yomdin, The geometry of critical and near-critical values of differentiable mappings, Math. Ann. 264 (4) (1983) 495–515.

[Yom2] Y. Yomdin, The set of zeroes of an "almost polynomial" function, Proc. Amer. Math. Soc. 90 (4) (1984) 538–542.

[Yom3] Y. Yomdin, Global bounds for the Betti numbers of regular fibers of differentiable mappings, Topology 24 (2) (1985) 145–152.

[Yom4] Y. Yomdin, Beta-spread of sets in metric spaces and critical values of smooth functions, preprint, MPI Bonn, 1983.

[Yom5] Y. Yomdin, Metric properties of semialgebraic sets and mappings and their applications in smooth analysis. in: J.M. Aroca, T. Sahcez-Geralda, J.L. Vicente (Eds.), Proceedings of the Second International Conference on Algebraic Geometry, La Rabida, Spain, 1984, Travaux en Course, Hermann, Paris, 1987, pp. 165–183.

[Yom6] Y. Yomdin, On functions representable as a supremum of a family of smooth functions. II, SIAM J. Math. Anal. 17 (4) (1986) 961–969.

[Yom7] Y. Yomdin, On functions representable as a supremum of a family of smooth functions, SIAM J. Math. Anal. 14 (2) (1983) 239–246.

[Yom8] Y. Yomdin, Critical values and representations of functions by means of compositions, SIAM J. Math. Anal. 17 (1) (1986) 236–239.

[Yom9] Y. Yomdin, Applications of a generalized differential of Lipschitzian functions in some results of singularity theory, I and II, preprints, Beer-Sheva, 1980.

[Yom10] Y. Yomdin, Some results on finite determinacy and stability, not requiring the explicit use of smoothness. in: Singularities, Part 2 (Arcata, Calif., 1981), Proceedings of the Symposium on Pure Mathematics, Vol. 40, A.M.S., Providence, RI, 1983, pp. 667–674.

[Yom11] Y. Yomdin, On representability of convex functions as maxima of linear families, preprint, MPI Bonn, 1983.

[Yom12] Y. Yomdin, Maxima of smooth families, III: Morse–Sard Theorem, preprint, MPI Bonn, 1984.

[Yom13] Y. Yomdin, Volume growth and entropy, Israel J. Math. 57 (3) (1987) 285–300.

[Yom14] Y. Yomdin, C^k resolution of semialgebraic mappings, Israel J. Math. 57 (3) (1987) 301–317.

[Yom15] Y. Yomdin, Local complexity growth for iterations of real analytic mappings and semi-continuity moduli of the entropy, Ergodic Theory Dyn. Systems 11 (1991) 583–602.

[Yom16] Y. Yomdin, Approximational complexity of functions and entropy, Proceedings of the Israel Mathematical Union Conference (Tel Aviv, 1987), Tel Aviv University, Tel Aviv, 1987, pp. 1–9.

[Yom17] Y. Yomdin, Approximational complexity of functions, Geometric Aspects of Functional Analysis (1986/87), Lecture Notes in Mathematics, Vol. 1317, Springer, Berlin, New York, 1988, pp. 21–43.

[Yom18] Y. Yomdin, Complexity of functions: some questions, conjectures and results, J. Complexity 7 (1991) 70–96.

[Yom-Com] Y. Yomdin, G. Comte, Tame Geometry with Applications in Smooth Analysis, Lecture Notes in Mathematics, No. 1834, Springer-Verlag, Berlin, 2002, 180pp.

[Y-E-B-S] Y. Yomdin, Y. Elichai, S. Birman, E. Spiegel, Synthetic textures in MPEG, preprint, 2002.

[Zer] M. Zerner, Approximation par des variétés algébriques dans les espaces Hilbertiens, Stud. Math. t. 38 (1970) 413–435.